

Arbejdsrapport fra Miljøstyrelsen

Nr. 49 1997

Datakvalitet og statistisk analyse i livscyklusvurdering

Arbejdsrapport fra Miljøstyrelsen
Nr. 49 1997

**Datakvalitet og statistisk analyse
i livscyklusvurdering**

Bo P. Weidema, Instituttet for Produktudvikling
Annette K. Ersbøll, Kongelige Veterinær- og Landbohøjskole
Leif Hoffmann, COWI

Rapporten er udarbejdet med tilskud fra Rådet vedr.
genanvendelse og mindre forurenende teknologi.

Det skal bemærkes, at de fremsatte synspunkter ikke
nødvendigvis dækkes af Rådet eller Miljøstyrelsen.

Indhold

Forord 5

1. Resumé og sammenfatning 6
 - 1.1 Formål 6
 - 1.2 Videnindsamling 6
 - 1.3 Beskrivelse af datakvalitet 8
 - 1.4 Statistisk analyse 11
 - 1.5 Forslag til opfølgende projekter 14
2. Summary 15
 - 2.1 Purpose of the Study 15
 - 2.2 Collection of Information on Statistical Analysis and LCA 15
 - 2.3 Description of Data Quality 17
 - 2.4 Statistical Analyses 19
 - 2.5 Needs for Further Investigations 22
3. Introduktion 24
4. Videnindsamling 25
 - 4.1 Introduktion 25
 - 4.2 Livscyklusvurdering og usikkerheder 26
 - 4.2.1 Sammendrag 26
 - 4.2.2 Sammenfattende vurdering 30
 - 4.3 Usikkerheder på industri- og miljødata 30
 - 4.3.1 Sammendrag 30
 - 4.3.2 Sammenfattende vurdering 31
 - 4.4 Generelle statistiske metoder i relation til livscyklusvurdering 31
 - 4.4.1 Sammendrag 31
 - 4.4.2 Sammenfattende vurdering 35
 - 4.5 Opfattelse og betydning af usikkerheder 35
 - 4.5.1 Sammendrag 35
 - 4.5.2 Sammenfattende vurdering 36
 - 4.6 Sammenfatning af erfaringer fra videnindsamlingen 36
5. Datakvalitet 37
 - 5.1 Introduktion til beskrivelse af datakvalitet 37
 - 5.1.1 Data og datakvalitet 37
 - 5.1.2 Datakvalitetsmålsætning 38
 - 5.1.3 Indsamling af data 39
 - 5.1.4 Indikatorer for datakvalitet 40
 - 5.1.5 Modifikation af kvalitetsindikatorer 43
 - 5.2 Case 1: Energiforbrug til landbrugsafgrøder 45
 - 5.2.1 Beskrivelse af casen 45
 - 5.2.2 Beskrivelse af datakvalitet 47
 - 5.2.3 Samlet vurdering af datakvaliteten 61

5.3	Case 2: Emissioner fra bageri	64
5.3.1	Beskrivelse af casen	64
5.3.2	Beskrivelse af datakvalitet	65
5.3.3	Samlet vurdering af datakvaliteten	69
5.4	Case 3: Energiforbrug til nedkøling af varer	71
5.4.1	Beskrivelse af casen	71
5.4.2	Beskrivelse af datakvalitet	73
5.4.3	Samlet vurdering af datakvaliteten	84
5.5	Sammenfattende vurdering af metoden til beskrivelse af datakvaliteten	84
6.	Statistisk analyse af usikkerheder	87
6.1	Introduktion til statistisk analyse	87
6.1.1	Statistisk analyse	87
6.1.2	Anvendte edb-programmer	90
6.2	Case 1: Energiforbrug til landbrugsafgrøder	94
6.2.1	Statistisk analyse af data	95
6.2.2	Simulering af data med @RISK	99
6.2.3	Vurdering af resultaterne af den statistiske analyse af data	102
6.2.4	Anvendelse af statistisk analyse livscyklusvurdering	104
6.3	Case 2: Emissioner fra bageri	109
6.3.1	Statistisk analyse af data	109
6.3.2	Vurdering af resultaterne af den statistiske analyse af data	112
6.4	Case 3: Levetid for køleskabe	113
6.4.1	Statistisk analyse af data	113
6.4.2	Vurdering af resultaterne af den statistiske analyse af data	116
6.5	Sammenfattende vurdering af den statistiske analyse	117
7.	Internationalt seminar om statistik og usikkerheder i forbindelse med livscyklusvurderinger	119
8.	Anbefalinger til videre arbejde/projektoplæg	122
	Referencer	126

Forord

Nærværende projekt er gennemført for Rådet vedrørende genanvendelse og mindre forurenende teknologi i perioden februar 1995 til maj 1997 med Mariane T. Hounum, Miljøstyrelsen som projektansvarlig.

Projektet er udført af en projektgruppe bestående af:

Leif Hoffmann, projektleder, Enviroplan A/S; fra 1. september 1996
COWI AS

Bo P. Weidema, Instituttet for Produktudvikling
Annette K. Ersbøll, Institut for Matematik og Fysik, Den Kongelige Veterinær- og Landbohøjskole

Herudover har Carsten Lassen, COWI AS bidraget med simuleringer af usikkerheder for én af de bearbejdede cases, og Marianne Suhr Wesnæs, Instituttet for Produktudvikling har bidraget med beskrivelse af datakvalitet for én af de bearbejdede cases.

1. udkast af rapporten er blevet præsenteret og diskuteret i Miljøstyrelsens "Følgegruppe om metodeudvikling på livscyklusvurderings- og produktudviklingsområdet". Arbejdet er blevet fulgt af og diskuteret med en gruppe bestående af:

Kim Christiansen, Enviroplan A/S; fra 1. september 1996 COWI AS
Erik Hansen, COWI AS

Prof. Mats Rudemo, Institut for Matematik og Fysik, Den Kongelige Veterinær- og Landbohøjskole

Prof. Knut Conradsen, Institut for Matematisk Modellering, DTU

1. Resumé og sammenfatning

1.1 Formål

Projektet er et indledende projekt, som har til formål at:

1. beskrive og anvende metoder til karakterisering af datakvalitet
2. beskrive og anvende statistisk analyse i livscyklusvurderinger
3. pege på emner/områder, hvor der er behov for yderligere udvikling og/eller afprøvning af metoder (hovedprojekt)

Datakvalitet

Formålet er at beskrive og videreudvikle en metode til karakterisering af datakvalitet samt illustrere anvendelsen heraf på en række udvalgte cases. Casene er udvalgt med henblik på at illustrere karakterisering af datakvalitet af forskellige typer af datasæt: landbrugsdata med oprindelse i forskellige lande, industrielle data samt data for levetider af og energiforbrugssdata for køleskabe.

Statistisk analyse

Formålet er endvidere at beskrive metoder til anvendelse af statistisk analyse i livscyklusvurdering i form af eksakte beregninger/approssimationer og simulation. De to forskellige beregningsprincipper er illustreret og diskuteret i casen med landbrugsdata. Analyse af industrielle data og anvendelse af forskellige fordelinger til at skønne levetider for køleskabe er illustreret med de to øvrige cases.

Hovedprojekt

Erfaringerne fra de to beskrevne formål er herefter udmøntet i en række forslag til opfølgende projekter. Ideerne bør herefter prioriteres og koordineres med andre igangværende projekter.

Projektindhold

Nærværende projekt består af tre hovedelementer:

- litteratursøgning/videnindsamling
- videreudvikling af metode til beskrivelse af datakvalitet og afprøvning af metoden på tre udvalgte cases
- statistisk analyse af data i de tre cases; i den ene case anvendtes såvel Taylors approksimation som simulation

Afslutningsvis gives en række forslag til opfølgende arbejde i form af projektidéer.

1.2 Videnindsamling

Videnindsamlingen er foregået ved søgning i allerede tilgængelig litteratur samt i en række databaser og ved kontakt til en række personer, som arbejder med håndtering af store datamængder (forskellige former for miljødata) og statistisk analyse heraf. De anvendte databaser er: Agris, Chemabs, Compendex*Plus, Enviroline, Mathsci og Pollution. Der er blevet anvendt følgende termer: usikkerhed, statistisk analyse, følsomhed, variation, estimering og standardafvigelse kombineret med livscyklusvurdering, risikovurdering, kvalitetskontrol,

kvalitetsmanagement og kvalitetssikring samt industri og processer/-emissioner/energi. Den indsamlede litteratur er behandlet inden for fire områder:

- livscyklusvurdering og usikkerheder
- usikkerheder på industri- og miljødata
- generelle statistiske metoder i relation til livscyklusvurdering
- opfattelse/betydning af usikkerheder

LCA og usikkerheder

Usikkerheder i forbindelse med livscyklusvurderinger kan opdeles i tre typer: tekniske, metodiske og epistemologiske usikkerheder. Usikkerhederne kan opstå på forskellige trin i livscyklusvurderingen i forbindelse med procesdata, systemdata, ved fastsættelse af den funktionelle enhed, ved karakterisering og ved valuering.

Forskellige "tommelfingerregler" kan anvendes til sammenligning og vurdering af forskellige typer af industrielle data for herved at reducere behovet for komplicerede statistiske beregninger. Metoden kræver dog en omfattende dokumentation i form af statistisk behandling af eksisterende datasæt.

Principper fra den økonomiske teori kan tilpasses til analyse af data anvendt i livscyklusvurderinger i form af pålidelighedsanalyser, validitetsanalyser, dominansanalyser og marginalanalyser. Disse metoder involverer omfattende og komplekse beregninger, hvorfor de bedst anvendes i forbindelse med EDB-programmer til livscyklusvurderinger.

Industri- og miljødata

Videnindsamlingen har kun givet ringe udbytte mht. oplysninger om generelle usikkerheder på industridata. For den europæiske plastindustri foreligger en række gennemsnitstal for f.eks. råstofforbrug, energiforbrug og emissioner i form af intervaller, som udtrykker spredningen i de pågældende data som følge af forskellig praksis og driftsforhold i forskellige lande.

Statistiske metoder

De statistiske metoder, der anvendes, kan i principippet opdeles i metoder, der benytter simulering af usikkerheder, og metoder der benytter beregning af usikkerheder ved f.eks. Taylor approksimationer.

Simulering er beskrevet i forhold til risikovurderinger i forbindelse med helbredsundersøgelser. Inputparametrene tillægges en fordeling, og en responsfunktion bestemmes ud fra de relevante inputparametre. Usikkerheden på responsfunktionerne kan så bestemmes ved f.eks. Monte Carlo simuleringer.

Variansen på summer og differencer kan relativt enkelt bestemmes, især hvis der ikke er korrelation mellem de indgående led; ofte vil man antage, at der ikke er nogen korrelation. Ved bestemmelse af variansen på produktet eller ratioen mellem to eller flere led benyttes f.eks. Taylors approksimationer.

Opfattelse af usikkerheder

Ved forespørgsel om usikkerhederne på f.eks. industrielle data fås ofte udsagn som "usikkerhed betyder, at variablen med 95%

sikkerhed ligger inden for et intervallet $x - y$ ", "der er en usikkerhed på 10%", eller "der er en usikkerhed på 2 MJ". Det første udsagn udtrykker et konfidensinterval, det andet udsagn formodes at udtrykke at udtrykke en variationskoefficient, og det tredje udsagn formodes at udtrykke en absolut størrelse f.eks. en spredning.

I nærværende projekt er det valgt at opfatte opgivne intervaller som 95% konfidensintervaller og beregne middelværdien og standardafvigelsen ud fra den forudsætning (hvis ikke andet er opgivet).

1.3 Beskrivelse af datakvalitet

En metode til karakterisering af datakvalitet er beskrevet teoretisk, og ved hjælp af tre udvalgte cases er metoden afprøvet.

Metode

I livscyklusvurderinger anvendes betegnelsen *miljødata* om data i de undersøgte processer, *systemdata* om flowet af råstoffer, energi og produkter gennem de undersøgte processer samt *ydelsesdata* vedrørende den funktionelle produktenhed. Ved udførelse af en komplet livscyklusvurdering anvendes endvidere data til *karakterisering* og *vurdering* af miljøpåvirkninger.

Datakvaliteten kan beskrives ved hjælp af et datakvalitetsindeks, som beskriver følgende parametre:

- Pålideligheden af data
- Fuldständigheden af data
- Tidsmæssig korrelation af data
- Geografisk korrelation af data
- Teknologisk korrelation af data

Hver enkelt parameter tildeles en score mellem 1 (bedst) og 5 (dårligst). Datakvaliteten kan f.eks. udtrykkes som følger (2,1,1,2,4).

Datakvalitetsindikatorerne kan bruges til at revidere dataindsamlingsstrategien og til at forbedre vurderingen af resultaternes pålidelighed. Datakvalitetsindikatoren for et givet datasæt (f.eks. anvendt mængde pesticid) kan angives i et skema med en kvalitativ beskrivelse af data-kvaliteten efterfulgt af den tildelte score. Data vil ofte være behæftet med en oprindelig usikkerhed udtrykt som variationskoefficient (f.eks. 2 kg/ha; CV: 5,5%, som f.eks. kan være en variation mellem forskellige år) og et datakvalitets indeks (2,1,1,2,4). Et eksempel herpå er vist i tabel 1.2.1.

Tabel 1.2.1

Datakvalitets-indikator for anvendt mængde pesticid med supplerende usikkerhed som følge af dårlig score på én eller flere datakvalitetsindikatorer.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed ¹
Oprindelig usikkerhed	Variation mellem år		CV: 5,5%
Pålidelighed	Skøn for værst tænkelige tilfælde af forbrug i private haver og lagerforsyning over år.	2	CV: 2,5%, estimat
Fuldstændighed	-	1	-
Tidsmæssig korrelation	-	1	-
Geografisk korrelation	Estimeret variation mellem bedrifter for samme afgrøde	2	CV: 17%, estimat
Teknologisk korrelation	Estimeret variation mellem afgrøder på samme bedrift.	4	CV: 33%, estimat
Samlet usikkerhed			CV: 38%

CV: Variationskoefficient.

Supplerende usikkerhed kan bestemmes ved at skønne, hvad en dårlig score (f.eks. pålideligheden) bidrager med af supplerende usikkerhed til datasættets samlede usikkerhed. Derved fås et datasæt med en større usikkerhed, men med en bedre datakvalitetsindikator (f.eks. 2 kg/ha; CV: 38% og (m,1,1,m,m)), hvor "m" fortæller, at usikkerheden er modificeret med et bidrag fra den pågældende indikator.

Case 1

Den første case tager udgangspunkt i energiforbrug til produktion af afgrøder til et skiveskåret rugbrød. Hertil anvendes rug, sukkerroer, hvede, hørfrø fra UK og fra Rusland, palmeolie, sesamfrø og solsikkefrø. Energiforbruget til produktion af disse afgrøder er beregnet ud fra energiforbrug til produktion af pesticider, N-, P- og K-gødning, kalk, husdyrgødning, markarbejde og tørring af afgrøder.

Af den sammenfattende tabel 5.2.25 ses, at de samlede usikkerheder varierer fra under 1% til 104%, og at de er domineret af enkelte meget store usikkerheder.

Denne observation kunne benyttes til at foreslå en "hurtig" metode til estimation af de samlede usikkerheder, idet mindre usikkerheder kan ignoreres. Datakvalitetsindeksene vil først og fremmest kunne anvendes som "check-lister" til at identificere de væsentligste usikkerheder for de enkelte data.

Case 2

Den anden case tager udgangspunkt i selve produktionen af rugbrødet, og emissioner fra bageriet af total-N og BOD til vand. Casen illustrerer et eksempel på karakterisering af industrielle data.

Beskrivelsen af datakvaliteten viser, at med en relativ god datakvalitet kan sammensætningen indskrænkes til at omfatte de data, der har væsentlig indflydelse på den samlede usikkerhed. I tilfælde, hvor datakvaliteten er ringere, kræves der en mere detaljeret anvendelse af datakvalitetsindikatorerne.

Case 3

Den tredje case tager udgangspunkt i energiforbruget til nedkøling af varer i private husholdninger, og levetiden af køleskabe indgår som en væsentlig parameter, idet energieffektiviteten for køleskabe er forbedret år for år. Der er anvendt statistiske oplysninger om salg af køleskabe i årene 1975 - 91. Eksemplet giver ligeledes mulighed for at illustrere anvendelse af informationer for tilsvarende produkter (andre køleskabstyper) og problematikken omkring allokering.

Sammenfatning

Casen om energiforbrug inden for landbrug viser, at de korrigerede variationskoefficienter domineres af enkelte store usikkerheder. Ud fra denne observation synes det rimeligt at foreslå, at datakvalitetsindeksene primært benyttes som checkliste for identificering af væsentlige usikkerheder, hvorefter der regnes videre med de væsentlige usikkerheder. Usikkerheden på det samlede resultat vil blive lettere undervurderet, hvilket kan imødegås ved introduktion af erfaringsbaserede korrektionsfaktorer. Uafhængigt af de supplerende usikkerheder kan der i visse tilfælde være tale om at justere middelværdierne som følge af viden om effektivisering af f.eks. energianvendelse til industrielle processer.

Anvendelse af den beskrevne metode til beskrivelse af datakvaliteten, kan synes overvældende. I praksis vil afsnittene "beskrivelse af casen" og "samlet vurdering af datakvaliteten" med den sammenfattende tabel kunne indgå i en rapport, mens det mellemliggende afsnit "beskrivelse af datakvaliteten" kunne placeres i bilag eller udelades. Alt efter de aktuelle datas kvalitet vil der være behov for en mere eller mindre detaljeret sammenfattende tabel, som det fremgår af casen om emissioner fra et bageri. Hvis datakvaliteten er dårlig, vil der være behov for at indikere, hvilke data der giver anledning til væsentlige supplerende usikkerheder med henblik på at forbedre datagrundlaget på netop disse punkter.

Detaljeringsniveauet bør ligeledes fastsættes i overensstemmelse med formålet med den aktuelle undersøgelse herunder hvem datakvaliteten skal dokumenteres overfor.

Under arbejdet med beskrivelse af datakvaliteten har det vist sig, at de anvendte datakvalitetsindikatorer ikke i alle tilfælde entydigt beskriver data i de bearbejdede cases, ligesom tildelingen af en score til de forskellige indikatorer ikke i alle tilfælde er entydig.

Tildeling af supplerende usikkerheder til de angivne datakvalitetsindex eller skøn over usikkerheder på de oprindelige data kan ligeledes være vanskeligt. I visse tilfælde kan lignende undersøgelser anvendes til at estimere usikkerheden på oprindelige data med en passende grundlse, f.eks. vil variationen i energiforbruget til et givet køleskab kunne anvendes som skøn over variationen i energiforbruget for et gennemsnitskøleskab.

1.4 Statistisk analyse

Den statistiske analyse af data er udført ved hjælp af eksakte metoder kombineret med Taylors approksimationer (SAS) og simuleringer i såvel SAS som @RISK (Monte Carlo simuleringer).

Metode

Ved beregning af middelværdi og varians for enkle funktioner, f.eks. summer eller differencer, er der anvendt eksakte metoder, mens beregningen af middelværdi og varians for sammensatte funktioner, f.eks. produkt eller ratio af to eller flere variabler, hvor eksakte udtryk kun vanskeligt kan bestemmes, er foretaget ved hjælp af Taylors approksimationer. Resultaterne ved anvendelse af Taylors approksimationer er uafhængig af typen af fordeling for de enkelte variable. Nøjagtigheden af Taylors approksimationer afhænger af størrelsen af variationen på de eksakte variabler i den sammensatte funktion. Nøjagtigheden er størst, når variansen på de enkelte variabler er lille. Beregninger udføres f.eks. ved programmering i SAS.

Middelværdi og varians for sammensatte funktioner (f.eks. produktet XY) bestemmes i simulering ved at trække et tilfældigt tal fra hver af variablernes respektive fordelinger. De udtrukne tal sammensættes til et tal i henhold til udtrykket for den sammensatte funktion. Udtagningen fra de respektive fordelinger og sammenregning til et tal gentages et antal gange (f.eks. 1000 gange). Middelværdi og varians for den sammensatte funktion bestemmes herefter som hhv. gennemsnit og varians af de gentagne udtrækninger og sammenregnede tal. Simuleringerne udføres f.eks. ved programmering i SAS eller @RISK.

Case 1

De to forskellige metoder til bestemmelse af middelværdi og varians er afprøvet på den tidligere beskrevne case om afgrøder til et rugbrød. Resultaterne af beregningerne er sammenfattet i tabel 1.3.1.

Tabel 1.3.1

Samlet oversigt over det totale energiforbrug og standardafvigelsen bestemt ved de 4 forskellige metoder.

Metode	Fordelingsantagelser	Energiforbrug MJ/kg	Standardafvigelse MJ/kg
Taylors approksimation	ingen	1,94	0,30
Simulation	normalfordeling	2,05	0,41
Simulation	lognormalfordeling	2,03	0,32
Simulation (@RISK)	normalfordeling	2,05	0,43

Ved anvendelse af Taylors approksimationer fås en middelværdi på 1,94 MJ/kg, som ligger lidt under middelværdien fundet ved simuleringerne, der resulterede i en middelværdi på henholdsvis 2,03 og 2,05 MJ/kg ved antagelse om lognormal og normal fordeling af inputvariablene. Ved anvendelse af Taylors approksimationer og simulering med antagelse om lognormalfordeling af inputvariablene fås en standardafvigelse på 0,30 og 0,32 MJ/kg, men simulering med antagelse om normalfordeling resulterer i en standardafvigelse på 0,41 og 0,43 MJ/kg.

Fejlanalyse

Fejlanalyse er gennemført ved hjælp @RISK, og den viser, at standardafvigelsen på udbyttet af rug er den usikkerhed, der har størst indflydelse på standardafvigelsen på det endelige resultat. Gennemregning af et eksempel, hvor standardafvigelsen på udbyttet af rug er sat lig nul, resulterer i en reduktion af standardafvigelsen på outputvariablen fra 0,43 til 0,21 MJ/kg.

Eliminering af usikkerheden på de øvrige inputvariabler vil kun resultere i mindre reduktioner af standardafvigelsen på outvariablen.

Cut-off

Der er skitseret to metoder til *cut-off*:

- udpegnig af processer, der bidrager uvæsentligt - med et samlet bidrag mindre end én standardafvigelse - til den samlede middelværdi
- udpegnig af inputvariabler, hvor en forbedring (reduktion) af standardafvigelsen ikke har nogen indflydelse på standardafvigelsen på den samlede middelværdi

For begge de skitserede metoder gælder, at der er behov for yderligere beskrivelse, diskussion og afprøvning af metoden.

Sammenligning

Der er skitseret to metoder til sammenligning af et produkt med et "miljømæssigt forbedret" produkt. Den ene metode sammenligner energiforbruget for den samlede livscyklus, mens den anden sammenligner energiforbruget for den del af de to produkters livscyklus der er forskellig. Middelværdierne er i begge tilfælde sammenlignet ved hjælp af konfidensintervaller. Resultaterne viser, at det i den valgte case ikke er signifikant forskel mellem det samlede energiforbrug for de to produkter, mens sammenligningen af energiforbruget til de dele, der er forskellige, viser en signifikant forskel.

Case 2

I den anden case bestemmes udledningen af BOD og total-N fra et bageri. Fordelingen mellem to udløb er ikke kendt, hvorfør emissioerne er beregnet for en lige fordeling mellem de to udløb og for en "worst case", hvor 90% af vandmængden udledes med de højeste koncentrationer af BOD og total-N. Resultaterne er sammenfatte i tabel 1.3.2.

Tabel 1.3.2
Udledning af BOD og total-N pr. kg brød.

Udledning	Forhold Nord-Øst	Estimat mg udledt/kg brød	Standardafvigelse mg/kg	Variationskoefficient %
BOD	1:1	1.30	0,0579	4,45
BOD	9:1	2.26	0.1039	4,60
Total-N	1:1	0.040	0.0050	12,50
Total-N	9:1	0.068	0.0086	12,65
BOD ¹	1:1	1.30	0.4836	37,20
Total-N ¹	1:1	0.040	0.0150	37,50

1. Usikkerheden omkring fordelingen af spildevandsmængde mellem det nordlige og det østlige udløb er der taget højde for ved at korrigere den samlede usikkerhed på BOD og total-N målingerne.

Variationskoefficienten ses at være uafhængig af det valgte forhold mellem udledninger til henholdsvis nord og øst. Som alternativ til "worst case" analysen er det undersøgt, hvad det betyder for den samlede standardafvigelse, hvis der korrigeres for usikkerheden omkring fordelingen af udledningen (henholdsvis nord og øst) ved at korrigere den samlede usikkerhed for BOD og total-N målingerne. Det er ikke muligt ud fra de foreliggende resultater at pege på den éne metode som den rigtige.

Case 3

Med tredje case demonstreres betydningen af valget af fordeling for levetiden for køleskabene. Levetiden er beskrevet ved enten en normalfordeling ($X \in N(\mu, \sigma^2)$) med middelværdien 14 år og en varians på 12,25 år² eller en gammafordeling ($X \in G(k, \beta)$) med en k-værdi på 16 samt en β -værdi på 0,875. Det samlede energiforbrug beregnet ud fra de forskellige fordelinger en vist i tabel 1.3.3.

Tabel 1.3.3

Energiforbrug (MWh) for køleskabe i 1992 ved anvendelse af forskellige antagelser om fordeling af antal køleskabe i funktion i 1992.

Fordeling	Energiforbrug MWh	Standardafvigelse MWh	Variationskoefficient %
N(14;3,5 ²)	770.531	79,01	0,0103
G(16;0,875)	770.058	79,02	0,0103

Det to valgte fordelinger resulterer i det samme totale energiforbrug og den samme variationskoefficient. En afbildning af aldersfordelingen af køleskabe på grundlag i de to valgte fordelinger viser ligeledes ens forløb af kurverne.

Sammenfatning

Anvendelse af statistisk analyse på de udvalgte cases viser, at Taylors approksimationer resulterer i et lavere samlet energiforbrug end simulation med såvel SAS som @RISK. Antagelse om lognormalfordeling af inputvariablerne resulterer i samme standardafvigelse som ved anvendelse af Taylors approksimationer, mens antagelse om normalfordeling, hvor der er risiko for udtrækning af negative værdier ved simuleringen, giver en lidt højere standard afvigelse.

Anvendelse af Taylor approksimationer og simuleringer (@RISK og SAS) til bestemmelse af den samlede usikkerhed på outputdata fra en udvalgt del af en livscyklusvurdering peger på, at Taylor approksimationer er håndterbart for mindre datasæt, mens simuleringer er bedre til at håndtere større datasæt. Fejlanalyse mv., som kræver simulering, udføres bedst med f.eks. @RISK, idet simulering med Taylors approksimationer kræver et mere omfattende programeringsarbejde.

1.5 Forslag til opfølgende projekter

Den gennemførte afprøvning af metoden til beskrivelse af datakvalitet og beregning af usikkerheder på resultatet af dele af en livscyklusvurdering har ledt til følgende forslag til videre arbejde:

- Afprøvning af metoden til beskrivelse af datakvalitet i en gruppe af brugere - scoring af datakvalitetsindikatorer og skøn/bestemmelse af supplerende usikkerheder betinget af datakvaliteten med henblik på at opnå konsensus om metoden
- Afprøvning af metoden på en række cases udvalgt til belysning af forskellige aspekter
- Udarbejdelse af paradigma for beskrivelse af datakvalitet på forskellige niveauer
- Test af kommersielt software til statistisk behandling af usikkerheder i forbindelse med livscyklusvurdering sammen med yderligere afprøvning af såvel @RISK som SAS
- Indsamling af industrielle data inden for udvalgte sektorer til brug for bestemmelse af generelle usikkerheder på industrielle data
- Usikkerheder på vurderingsdata forekommer, og betydningen heraf bør undersøges nærmere.
- Korrelationer mellem inputvariabler kan påvirke outputvariablen, hvis ikke størrelsen af korrelationerne kendes og medtages i beregningerne.
- *Cut-off* og fejlanalyse er kort berørt, men der er behov for yderligere beskrivelse af og test af metoderne med henblik på at opnå konsensus om, hvordan metoderne kan anvendes i livscyklusvurdering.
- Generel vejledning i anvendelse af usikkerhedsvurderinger i livscyklusvurdering. Vejledningen skal beskrive forudsætninger, metoder samt dokumentere de foreliggende metoder.

2. Summary

2.1 Purpose of the Study

The project is a preliminary project with the purpose of:

- describe and use methods to characterise data quality
- describe and use statistical analyses in life cycle assessment
- identify areas with needs for further development and/or testing of methods

The purpose is to describe and develop a method of characterising data quality and illustrate the use of the method in selected cases. The cases are selected in order to illustrate the characterisation of the data quality of different kinds of data sets: agricultural data from different countries, industrial data and data concerning life time of refrigerators.

The purpose is to describe methods of statistical analysis in life cycle assessment by using exact calculations/approximations and simulations. The different calculation principles are illustrated and discussed in the agriculture case. Analyses of industrial data and use of different distributions to describe life time are illustrated in the two other cases.

Experience from the above mentioned purposes is afterwards expressed in a number of recommendations for further investigations. The different issues have to be coordinated with already initiated studies.

This project includes:

- systematic collection of information on statistical analyses and life cycle assessment
- further development of a method to characterise data quality and test of the method on three selected cases
- statistical analysis of data from three cases; in one case Taylor approximation has been used as well as simulation

Finally, suggestions for further investigations are presented.

2.2 Collection of Information on Statistical Analysis and LCA

The collection of information has been done by searching in available literature, in databases and by contact to people working with handling of high amounts of data and statistical analyses. The collected literature has been described under the headlines:

- life cycle assessment and environmental data

- uncertainties on industrial and environmental data
- general statistical methods in relation to life cycle assessment
- understanding/interpretation of uncertainties

LCA and uncertainties

Uncertainties in life cycle assessment can be divided into three types: technical, methodological and epistemological uncertainties. The uncertainties can be introduced in different steps in the life cycle in relation to process data, system data, unit data, characterisation data and valuation data.

Different "roles of thumb" can be used in comparison and assessment of industrial data and hereby reduce the needs for complicated statistical calculations. The methodology requires comprehensive documentation like eg statistical assessment of existing data sets.

Principles from economic theory can be adapted in order to analyse data in life cycle assessment by using reliability analysis, validity analysis, dominance analysis and marginal analysis. These methods involve comprehensive and complex calculations and therefore, the need for automated computation.

Industrial and environmental data

The collection of information has only resulted in data from one source: PWMI (European Centre for Plastics in the Environment), who has published data from the European plastic industry. Average data are available for eg consumption of raw material, energy consumption and different emissions. The gross energy consumption is given as a range, expressing the spread in the data as a consequence of different operation conditions in different countries.

Statistical methods

The statistical methods can be divided in methods based on simulations and methods using exact calculations and approximations.

Simulations are described in relation to risk assessment in connection with public health investigations. The input parameters are described by statistical distributions and a response function can be determined by using relevant input parameters. The uncertainty on the response function (the output variable) can be determined by eg "Monte Carlo" simulations.

The variation of sums and differences can be determined by relative simple formulas especially if the variable is not correlated. Calculation of variation of products or ratios of two or more variables can be done by eg Taylor approximations.

Understanding of uncertainties

Uncertainties on industrial data is often expressed as: "uncertainty means the variable is in the range with a confidence of 95%", "the uncertainty is 10%" or "the uncertainty is 2 MJ". The first statement is supposed to express confidence limits, the second statement is supposed to express a coefficient interval, and the third statement is supposed to express an absolute number, eg a spread.

In this project, intervals have been interpreted as 95% confidence intervals, and the mean and spread has been calculated by using this assumption.

2.3 Description of Data Quality

A method to characterise data quality is described theoretically, and the method is tested in three cases.

Methodology

In life cycle assessment, the term environmental data expresses data from the considered process, system data expresses data on the flow of raw materials, energy and products, and unit data expresses information on the functional unit. In a complete life cycle assessment also characterisation and valuation data are included.

The data quality can be described by a data quality index describing the following parameters:

- reliability of data
- completeness of data
- time-related correlation
- geographical correlation
- technological correlation

The individual parameters are assigned with a score between 1 (the best) and 5 (the worst). The data quality can be expressed as follows: (2,1,1,2,4).

The data quality index can be used in revision of the data collection strategy and to improve the assessment of the reliability of the data. The data quality indicator for a set of data can be presented in a table along with a qualitative description of the data quality followed by the given score. Data will often have an uncertainty expressed as a coefficient of variance (eg 2 kg/ha; CV: 5.5% - variation between years) and a data quality index (2,1,1,2,4) as shown in table 2.2.1.

Table 2.2.1
Data quality indicator for used amount of pesticides with supplementary uncertainties.

Data quality indicator	Data quality	Score	Uncertainty ¹
Uncertainty	Variation between years.		CV: 5.5%
Reliability	Worst case estimate for use of pesticides in private gardens an stock variations between years.	2	CV: 2.5%, estimate
Completeness	-	1	-
Time-related correlation	-	1	-
Geographical correlation	Estimated variation between different farms, same crop.	2	CV: 17%, estimate
Technological correlation	Estimated variation between different crops, same farm.	4	CV: 33%, estimate
Total uncertainty			CV: 38%

CV: Coefficient of variance.

The supplementary uncertainty can be determined by estimating the influence of a low score to the total uncertainty of the data set. The data set thereafter has a bigger uncertainty but a better data quality indicator (eg 2 kg/ha; CV: 38% and (m,1,1,m,m)) where "m" denotes that the uncertainty has been modified with a contribution from that specific indicator.

Case 1

The first case is based on energy consumption to production of crops to production of a rye bread. The crops are rye, sugar beets, and wheat from Denmark, linseed from UK and Russia, palm fruit, sesame seed and sunflower seed. The energy consumption to the production of bread is calculated from the energy consumption to production of pesticides, N-, P-, K-fertilizer, CaCO₃, crop drying, field operations, and manure.

The total uncertainties varies from below 1% to 104% and are dominated by single very high uncertainties.

This observation can be used to propose a quick method to estimate the total uncertainty as the small uncertainties then will be ignored. The data quality index can also be used to identify the most important uncertainties.

Case 2

The second case is based on production of rye bread at a bakery and emissions of total-N and BOD with sewage water. This case is an example on characterisation of industrial data. The description of the data quality has shown that the summary of the characterisation of data sets with a relative good quality can be reduced to focus on the data with essential influence on the total data quality. In cases where the data quality is not as good, a more detailed use of the data quality indicators is necessary.

Case 3

The third case is based on energy consumption to cool down food in private households. The life time for refrigerators is an important parameter as energy efficiency for refrigerators has been improved

from year to year. Statistical information on the sale of refrigerators in the years 1975 - 91 have been used in the estimation of the distribution of the age of refrigerators in function in 1992. The example includes use of information from corresponding products and allocation problems.

Summary

The case involving agricultural data shows corrugated coefficient of variance varying from 1% to 104% dominated by single uncertainties. From this observation, it seems obvious to propose that data quality indices primarily is used as a "check-list" to identify dominating uncertainties. These dominating uncertainties are supposed to be used in the following calculations. The uncertainty on the final result will be slightly under-estimated, which can be corrugated by introducing knowledge based correction factors. Adjustment of the means as a consequence of knowledge on eg effectivity of energy consumption in industrial processes can be a possibility.

The described method of characterising data quality might seem overwhelming. In praxis "Description of the case" and "Assessment of the data quality" can be included in the reporting and "Description of the data quality" can be placed in annex or omitted. The summarising can be more or less detailed depending on the actual data quality as shown in the case involving industrial data. If the quality of the data is low, there is a need for indicating data contributing with essential supplementary uncertainties with the purpose of improving the data for these areas.

The level of details is recommended to be defined in agreement with the goal of the specific study and who the data quality has to be documented to.

When working with the characterising of the data quality, it has become clear that the data quality indicators not in all cases can describe the specific data in the concerned cases as well as assigning a score for the different indicators cannot be done unequivocal in all cases.

Assigning of supplementary uncertainties to the data quality index or estimates of uncertainties on the original data can also be difficult. In some cases similar investigations can be used in the estimation of uncertainties on original data with an appropriate argumentation eg use of variation on the energy consumption of a specific refrigerator as an estimate of the variation on the energy consumption of an average refrigerator.

2.4 Statistical Analyses

The statistical analyses of data are done by using exact methods combined with Taylor approximations (SAS) and by simulations using SAS as well as @RISK (eg Monte Carlo simulations).

Method

Calculation of mean and variation for simple functions is done by exact methods whereas calculation of mean and variation for

combined functions is done by using Taylor approximations. The results of the exact and the Taylor approximations are independent of the distribution of the input variables. The precision of the Taylor approximations depend of the magnitude of the variation on the specific variables in the combined functions ie the precision is highest when the variation on the input variable is low. These calculations can be done by programming in SAS.

Mean and variation of combined functions can be estimated by simulation. A random number is selected for each variable and the output variable is determined. This process is repeated eg 1000 times. Mean and variation can be determined on the different output variables. These simulations can be done by programming in SAS or @RISK.

Case 1

The different methods to determine mean and variation have been tested on the described case based on energy consumption to production of crops to rye bread. The results are summarised in table 2.3.1.

Table 2.3.1
Overview of the total energy consumption (mean) and spread determined by different methods.

Method	Assumption on distribution	Energy consumption MJ/kg	Spread MJ/kg
Taylors approximation	none	1.94	0.30
Simulation (SAS)	normal distribution	2.05	0.41
Simulation (SAS)	log normal distribution	2.03	0.32
Simulation (@RISK)	normal distribution	2.05	0.43

The Taylor approximations result in a mean at 1.94 MJ/kg that is below the mean found with simulations. The Taylor approximations and the simulation assuming log normal distribution of the input variables result in a spread at 0.30 and 0.32 MJ/kg respectively, whereas the other simulation results in a spread at 0.41 and 0.43 MJ/kg.

Error analysis

Error analysis is made by use of @RISK, and the results show that the spread at the yield of rye is the uncertainty with the highest influence on the total uncertainty. If the spread on the yield of rye is assumed to be zero the total uncertainty is reduced from 0.43 to 0.21 MJ/kg.

Illumination of the uncertainty on the other input variables will only result in minor reductions in the total uncertainty.

Cut-off

Two methods to look at *cut off* have been presented:

- pointing out processes contributing insignificant to the final energy consumption with a total value below one spread

- pointing out input variables where a reduction of the spread will have no influence on the total uncertainty

For both methods, it is necessary to describe, discuss and test the method further.

Comparison of two products Two methods to compare a product with an "environmental improved" product have been described. The first method compares the total energy consumption whereas the other method compares the energy consumption for the part of the life cycle that differs from each other. The mean are in both methods compared by use of means and confidens intervals. The results show that the improvement is not significant if considering the total energy consumption, whereas the part of the life cycle that differs from each other is significant different.

Case 2

In the second case, the emission of BOD and total-N from a bakery is considered. The distribution of sewage water from two outlets are not known. The emission is therefore calculated for *worst case* where 90% of the sewage water has the highest concentration of pollutant. The results are summarised in table 2.3.2.

Table 2.3.2
Outlet of BOD and total-N pr. kg bread.

Outlet	Ratio North-East	Estimate mg /kg bread	Spread mg/kg	Coefficient of variation %
BOD	1:1	1.30	0.0579	4.45
BOD	9:1	2.26	0.1039	4.60
Total-N	1:1	0.040	0.0050	12.50
Total-N	9:1	0.068	0.0086	12.65
BOD ¹	1:1	1.30	0.4836	37.20
Total-N ¹	1:1	0.040	0.0150	37.50

1. The uncertainty on the distribution of the outlet between north and east has been corrugated by the supplementary uncertainty assigned to the original uncertainty.

Case 3

The third case demonstrates the importance of the choice of distribution to describe the life time for refrigerators. The life time has been described by either a normal distribution ($X \in N(\mu, \sigma^2)$) with a mean at 14 year and a variation at 12.25 year² or a gamma distribution ($X \in G(k, \beta)$) with a k-value at 16 and a β -value at 0.875. The total energy consumption calculated by use of the different distributions is shown in table 2.3.3.

Table 2.3.3

Energy consumption (MWh) for refrigerators in 1992 calculated under assumption of different distributions of the age of refrigerators.

Distribution	Energy consumption MWh	Spread MWh	Coefficient of variation %
N(14;3.5 ²)	770.531	79.01	0.0103
G(16;0.875)	770.058	79.02	0.0103

The two selected distributions result in the same total energy consumption and the same coefficient of variation. A plot of the age distribution shows similar curves.

Summary

Statistical analysis used at the three selected cases shows that Taylor approximations result in a lower total energy consumption than simulation with @RISK as well as SAS. Assumption of log normal distribution of the input variables results in the same spread as Taylor approximations whereas assumption of normal distribution (with a risk of selecting negative values) results in a higher spread.

The use of Taylor approximations and simulations in the determination of the uncertainty on output variables from selected parts of a life cycle point out that Taylor approximations are useable for small data sets whereas simulations by eg @ is better to handle bigger data sets. Error analyses that require simulations are done better by use of @RISK as simulation combined with Taylor approximations requires extensive programming.

2.5 Needs for Further Investigations

Test of the method to characterise data quality and calculation of uncertainties of the final result of parts of a life cycle assessment has identified the following areas with needs for further development of method and/or testing of methods:

- Test of the method to describe data quality by a group of users with the purpose of obtaining consensus about the method.
- Test of the method on additional cases with the purpose of clarifying different aspects.
- Preparation of a paradigm for description of data quality on different levels.
- Test of commercial software to statistical treatment of uncertainties in relation to life cycle assessment.
- Compiling of industrial data from selected industrial sectors to be used in a general estimate of uncertainties on industrial data.

- Description and assessment of the uncertainties on weighting data.
- Investigation of correlations and the influence hereof.
- Definition of error analysis and "cut off" in relation to life cycle assessment and test on selected cases.
- Preparation of a general guideline about uncertainty analysis in life cycle assessment.

3. Introduktion

Litteratursøgning

Beskrivelse af datakvalitet og statistisk analyse har indtil ideén til projektet opstod kun haft ringe udbredelse inden for livscyklusvurdering. Dette ses også af det ringe udbytte af litteratursøgningen (kap. 4).

Datakvalitet

De indledende forsøg på at beskrive datakvaliteten i relation til livscyklusvurdering er gjort i Weidema (1994), som inspireret af Funtowicz & Ravetz (1990) udviklede en datakvalitetsmatriks til beskrivelse af datas pålidelighed, fuldstændighed og tidsmæssige, geografiske og teknologiske korrelation. Hver datakvalitetsindikator kan gives en værdi mellem 1 og 5.

Metoden er videreudviklet i nærværende projekt og er beskrevet i detaljer i afsnit 5.1; i afsnittene 5.2, 5.3 og 5.4 er metoden afprøvet på udvalgte cases.

Usikkerheder i livscyklusvurdering

I forbindelse med det nordiske LCA projekt (Lindfors *et al.*, 1995) arbejdede projektgruppen med en præcisering og definition af usikkerheder og statistiske begreber i forbindelse med LCA (Hoffmann *et al.*, 1995) inspireret af bl.a. Weidema (1993), som opregner tre typer af usikkerheder i forbindelse med en livscyklusvurdering: tekniske usikkerheder, metodiske usikkerheder og epistemologiske usikkerheder. Usikkerhederne relateret til forskellige trin i en livscyklusvurdering kan fremkomme ved anvendelse af procesdata, systemdata, fastsættelse af den funktionelle enhed, ved karakterisering og valuering. Ved hjælp af statistiske metoder illustreres anvendelse af statistisk analyse i livscyklusvurdering i et simpelt konstrueret eksempel. Tilsvarende statistisk analyse af de tre cases er beskrevet i kapitel 6.

Forprojekt

Under præsentation af de opnåede resultater på et seminar afholdt i forbindelse med det nordiske LCA-projekt, viste der sig stor interesse for det indledende arbejde. Idéen om et opfølgende projekt blev forelagt Miljøstyrelsen, og et forprojekt blev formuleret sammen med en skitse til et hovedprojekt. Ideer til videre arbejde formuleres som en del af nærværende projekt (kap. 8).

Statistiske forudsætninger

I nærværende rapport angives en variabel som en middelværdi og en standardafvigelse, f.eks. middelværdi 10 MJ med en standardafvigelse på 2 MJ. Denne standardafvigelse kan også udtrykkes relativt i forhold til middelværdien eller som procentdel af middelværdi, hvorved fås variationskoefficienten, her 20%. I tilfælde, hvor data er opgivet som intervaller, forudsættes det at være 95% konfidensintervaller, som tilnærmedesvis kan bestemmes som $\mu \pm 2\sigma$. I tilfælde, hvor data er angivet som "worst case" estimater, antages estimatet at svare til 97,5% fraktilen i en normalfordeling.

4. Videnindsamling

4.1 Introduktion

I nærværende kapitel er resultatet af en række litteratursøgninger, foretaget i foråret 1995, beskrevet. Litteratursøgningerne er dels foretaget i internationale databaser dels ved skriftlig kontakt til eksperter inden for såvel livscyklusvurdering som statistisk analyse af kompleksse systemer. Endvidere er Danmarks Miljøundersøgelse og Miljøstyrelsen kontaktet med henblik på fremskaffelse af resultater af moniteringer gennem en længere periode.

Projektgruppen har derudover været medarrangører af et internationalt seminar om statistik og usikkerheder i forbindelse med livscyklusvurderinger. Det internationale seminar er beskrevet i kapitel 7.

Databaser

Ved litteratursøgningen er følgende databaser anvendt (ESA-IRS):

- Agris - landbrug
- Chemabs - kemi
- Compendex*Plus - teknik
- Enviroline - miljø
- Mathsci - matematik (efter 1985)
- Pollution - forurening, miljø

Søgeprofil

Ved litteratursøgningen er følgende statistiske termer anvendt:

- Usikkerhed
- Statistisk analyse
- Følsomhed
- Variation
- Estimation
- Standardafvigelse

Disse statistiske termer er kombineret med følgende grupper af søgeord:

- Livscyklusvurdering, risikovurdering, kvalitetskontrol, kvalitetsmanagement, kvalitetssikring
- Industri og processer
- Industri og emissioner
- Industri og energi

Endvidere er følgende sekvenser anvendt som søgeord:

- Vurdering af usikkerheder
- Bestemmelse af usikkerheder

De udvalgte referencer er beskrevet i fire afsnit, som omhandler:

- 1) Livscyklusvurdering og usikkerheder
- 2) Usikkerheder på industri- og miljødata
- 3) Generelle statistiske metoder i relation til livscyklusvurdering
- 4) Opfattelse og betydning af usikkerheder

Nogle af de udvalgte referencer er behandlet flere gange, idet de fx omhandler både livscyklusvurdering og usikkerheder samt generelle statistiske metoder i relation til livscyklusvurdering.

4.2 Livscyklusvurdering og usikkerheder

4.2.1 Sammendrag

Usikkerheder i forbindelse med livscyklusvurderinger er i litteraturen behandlet på følgende niveauer:

- Teoretiske betragtninger om usikkerheder (Hoffmann *et al.* (1995), Tillmann *et al.* (1994), van den Hout (1994), van Hees (1994), Weidema (1993), Young & Vanderburg (1994))
- Metodemæssige betragtninger om håndtering af usikkerheder (Franklin Associates LtD (1993), Heijungs (1994; 1995))
- Anvendelse af usikkerhedsvurderinger (Rønning *et al.* (1993))

Tillmann *et al.* (1994) og Young & Vanderburg (1994) beskriver problemer med usikkerheder i forbindelse med fortolkning af resultater af en livscyklusvurdering som følge af de betragtede systemers kompleksitet, samt usikkerheder ved vurdering af miljøeffekter. De giver dog ikke konkrete anvisninger på håndteringen af usikkerheder.

van Hees (1994) arbejder med anvendelse af livscyklusvurdering til produktudvikling (o.a.), hvor det netop er forskellen i resultaterne og usikkerheden på disse, der er relevant. Usikkerheden kan findes i forskellige forhold:

- 1) manglende data
- 2) målesikkerhed
- 3) forskelle mellem processer
- 4) gamle data
- 5) systemgrænser

Oplysninger anvendt i livscyklusvurderinger søges ofte i databaser, som influerer på punkterne 1 - 4, mens proceduren for livscyklusvurdering influerer på punkterne 1 - 5. Anvendelsen af databaser og valg af livscyklusvurderingsprocedure giver således tilsammen usikkerheder på det endelige resultat. Arbejdshypotesen er, "*at der findes et tilstrækkeligt antal statistisk uafhængige data, hvis varians vil kunne ophæve hinanden og derved give et præcist resultat*". Endvidere vil de anvendte data ofte være korrelerede eller korrelationer vil blive indført ved forsøget på at skønne manglende data.

van den Hout (1994) introducerer en mere systematisk opdeling af usikkerheder ved gennemførelse af livscyklusvurderinger:

- usikkerheder på de anvendte data
- usikkerheder som følge af metodiske simplifikationer
- usikkerheder i grænsefladen mellem videnskab og samfundet

Dette svarer til den opdeling, der præsenteres af Weidema (1993):

- tekniske usikkerheder - skyldes variation i de indsamlede data
- metodiske usikkerheder - skyldes antagelser mv. i forbindelse med definition af produktssystem og procedurer for dataindsamling
- epistemologiske usikkerheder - skyldes oversete problemstillinger, udeladelser eller mangel på repræsentativitet

Weidema (1993) opdeler anvendelsesformerne for statistisk analyse eller følsomhedsanalyse i forbindelse med livscyklusvurdering til at:

- udelukke processer mv. af mindre betydning (reducere dataindsamlingen)
- kvalificere de indsamlede data
- kvalificere det endelige resultat

Hoffmann *et al.* (1995) beskriver, at usikkerheder kan opstå på følgende trin i proceduren for gennemførelse af livscyklusvurdering i forbindelse med:

- procesdata
- systemdata
- fastsættelse af den funktionelle enhed
- karakterisering
- valuering

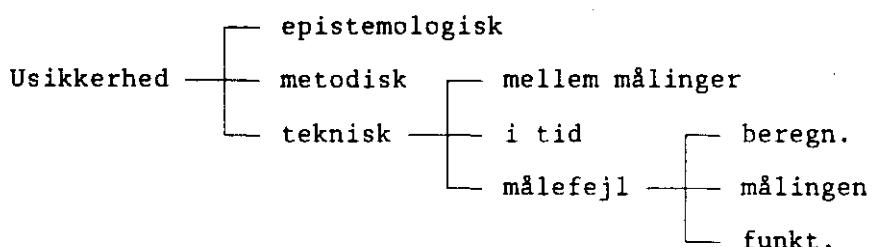
Usikkerhederne er nærmere forklaret i tabel 4.2.1.

Tabel 4.2.1

Usikkerheder relateret til de forskellige trin i en livscyklusvurdering (Hoffmann et al., 1995).

		Usikkerhed
1	Specifikation af input/output data for proces p pr. tidsenhed [fx kg CO ₂ /år/proces p]	U(proces)
2	Normalisering af input/output data til komponent x fra proces p [fx kg CO ₂ /kg komponent x]	
3	Normalisering af input/output data for komponent x til produkt P [fx kg CO ₂ /kg komponent x × kg komponent x/kg produkt P = kg CO ₂ /kg P]	U(system)
4	Summering af input/output data for produkt P fra forskellige komponenter [fx $\sum_{x=1}^n$ kg CO ₂ /kg P _x = kg CO ₂ /kg P]	
5	Normalisering af input/output data for produkt P til funktionel enhed FE (til at sammenligne A med B) [kg CO ₂ /FE _P]	U(enhed)
6	Klassificering [E ₁ , E ₂ , .. E _n → drivhuseffekt]	
7	Karakterisering [E _i → CO ₂ -eq.(E _i) → Σ CO ₂ -eq.]	U(karakterisering)
8	Valuering [Σ CO ₂ -eq. (potentiale for drivhuseffekt), Σ CFC-11-eq. (ozonreduktionspotentiale),]	U(valuering)

En mere detaljeret inddeling af usikkerheder er vist i figur 4.2.1.



Figur 4.2.1

Sammenhæng mellem forskellige usikkerheder (Hoffmann et al., 1995).

Statistiske metoder er anvendt på to konstruerede eksempler (luftemissioner ved energiproduktion på basis af affald og kul). Ved hjælp af konfidensintervaller er emissionerne i de to eksempler sammenlignet. Eksemplet viser, at nogle emissioner kan være signifikant forskellige, mens det for andre ikke kan påvises (Hoffmann et al., 1995).

Metodiske betragtninger

Franklin Associates Ltd (1993) har med baggrund i deres erfaringer givet tommelfingerregler for, hvornår forskellige typer af industrielle data er forskellige. Deres generelle opfattelse er, at de data, der anvendes i LCA, ikke opfylder de formelle krav til en statistisk analyse af data (tilfældig stikprøve, store populationer og normalfordeling). Med udgangspunkt i generelle statistiske betragtninger har de opstillet nogle grænser for, hvornår de opfatter data som forskellige. De statistiske betragtninger er dog ikke nærmere beskrevet. Forskellen i energiforbrug og affald efter brugsfasen (vægtbasis) betragtes som signifikant, hvis den er mindst 10%, mens forskellen i affald efter brugsfasen (volumenbasis), industrielt affald, emissioner til luft og vand betragtes som signifikant, hvis den er mindst 25%.

Heijungs (1994; 1995) går mere systematisk til værks. Heijungs (1994) introducerer vektor-/matrixregning i forbindelse med livscyklus kortlægning ud fra principper anvendt inden for økonomisk teori. Metoden giver mulighed for udførelse af:

- 1) *marginalanalyse*: undersøgelse af effekten af små ændringer på det samlede resultat
- 2) *"pålidelighedsanalyse"*: undersøgelse af pålideligheden af de indgående parametre
- 3) *dominansanalyse*: undersøgelse af hvilken proces, der giver dominerende bidrag til en bestemt emission.

Heijungs (1995) giver baggrunden for anvendelse af usikkerhedsanalyse i forbindelse med "hot spot" identifikation (livscyklus-screening), og illustrerer derefter beregningsmetoderne med et simpelt eksempel, som er baseret på en lineær udgave af fejlophobningsloven.

$$\Delta x = \left| \frac{\partial x}{\partial a} \right| \Delta a + \left| \frac{\partial x}{\partial b} \right| \Delta b + \left| \frac{\partial x}{\partial c} \right| \Delta c + \dots \quad (1)$$

Der gives ligeledes et eksempel på anvendelse af matrixregning til bestemmelse af usikkerheder med udgangspunkt i en tidligere beskrevet metode (Heijungs, 1994). Ved indføring af kendte eller skønnede usikkerheder på parametrene kan den samlede usikkerhed såvel som usikkerheden på de enkelte processer eller inputparametre beregnes.

Usikkerhedsvurderinger

Rønning *et al.* (1993) har under anvendelse af skønnede usikkerheder på 1%, 10%, 25%, 50% og 100% på forskellige parametre og forskellige faser i en livscyklus beregnet usikkerheden på de parametre, som sammenlignes. Usikkerhedsanalysen er foretaget på alle emissioner til luft og vand, affald, energiforbrug og transport i forbindelse med sammenligning af to malingsprodukter. Beregningerne er baseret på Gauss' fejlophobningslov

$$\Delta x = \sqrt{\left(\frac{\partial x}{\partial a} \Delta a \right)^2 + \left(\frac{\partial x}{\partial b} \Delta b \right)^2 + \left(\frac{\partial x}{\partial c} \Delta c \right)^2 + \dots} \quad (2)$$

Usikkerheder på de enkelte indgående parametre er skønnet, se tabel 4.2.2.

Tabel 4.2.2

*Skønnede usikkerheder på forskellige parametre (Rønning *et al.*, 1993).*

	Produktion af råmaterialer	Produktion af færdigvarer	Anvendelse af færdigvarer
Transport	100%	25%	50%
Energiforbrug	25%	25%	25%
Emissioner til luft	10%	10%	1%
Emissioner til vand	100%	10%	50%
Kemisk affald	100%	10%	50%
Fast affald	100%	10%	50%

Resultatet af sammenligningen er givet som f.eks. potentiale for drivhuseffekt (CO_2 -ækv./1.000 l): 862 ± 142 for vandbaseret maling henholdsvis 1.211 ± 104 for opløsningsmiddelbaseret maling.

4.2.2 Sammenfattende vurdering

Terminologi

Den mest stringente terminologi er præsenteret af Weidema (1993), som inddeler usikkerheder i tekniske, metodiske og epistemologiske usikkerheder og Hoffmann *et al.* (1995), der beskriver kilderne til usikkerheder i livscyklusvurdering som usikkerheder i forbindelse med procesdata, systemdata, fastsættelse af en funktionel enhed, karakterisering og valuering. van Hees (1994) og van den Hout (1994) giver ligeledes en beskrivelse af kilderne til og typer af usikkerheder, som mere eller mindre kan sammenlignes med ovennævnte, dvs. de bidrager ikke med nye vinkler i forhold til terminologien.

Behandling af usikkerheder

Usikkerheder er kun i ringe grad inddraget i hidtil udførte livscyklusvurderinger. Rønning *et al.* (1994) anvender skønnede usikkerheder på transport, energiforbrug, emissioner til vand og luft, kemisk og fast affald i forbindelse med produktion af råmaterialer, færdigvarer og ved anvendelse af færdigvarer. Det fremgår ikke, hvordan de forskellige usikkerheder er skønnet, men de er formodentligt baseret på erfaringer eller kvalificerede skøn; ved en systematisk indsamling af industrielle data vil de givne usikkerheder kunne be- eller afkraestes. Heijungs (1995) beskriver en metode til behandling af usikkerheder (kendte eller skønnede) i forbindelse med livscyklus-screening. Metoden inddrager matrixregning og er derfor ikke umiddelbar anvendelig ved "håndregning", men derimod anvendelig til indbygning i et data-program. Rønning *et al.* (1994) og Heijungs (1995) argumenterer for anvendelse af forskellige former for Gauss' fejlophobningslov.

Tommelfingerregler

Franklin Associates Ltd (1993) præsenterer retningslinier for, hvornår de vurderer to resultater som værende forskellige ud fra størrelsen af forskellen angivet i %; forskellen er angivet til 10% for f.eks. energiforbrug og 25% for f.eks. emissioner til luft. Metoden er i principippet den enkleste til anvendelse i forbindelse med livscyklusvurderinger, men den kræver en omfattende dokumentation i forbindelse med bestemmelse af størrelsen af "signifikante forskelle". Dokumentation af metoden er ikke offentlig tilgængelig.

4.3 Usikkerheder på industri- og miljødata

4.3.1 Sammendrag

APME (Association of Plastics Manufacturers in Europe)/PWMI (European Centre for Plastics in the Environment) har foranlediget udarbejdelsen af "økoprosfiler" for en række plastmaterialer (Boustead 1993a, b, c; 1994). Økoprofilerne er baseret på oplysninger indhentet fra en række europæiske virksomheder vha. spørgeskemaer, og de præsenterede oplysninger repræsenterer således en form for gennemsnitligt råstofforbrug, energiforbrug og emissioner, uden at der angives statistiske informationer. Det samlede energiforbrug (GER) er givet som vægtet gennemsnit med angivelse af interval (mindste og største energiforbrug) for en del af produkterne; se tabel 4.3.1. Inter-

vallet udtrykker ikke (alene) statistisk (tilfældig) variation, men derimod forskellig praksis og driftsforhold i forskellige lande.

Tabel 4.3.1

Samlet energiforbrug (GER) for en række forskellige mellemprodukter og forskellige plasttyper. Referencer i givet i noterne til tabellen.

Produkt	Antal producenter	Vægtet gennemsnit (GJ/ton)	Interval (GJ/ton)	Reference
Ethylen	19 ? ¹	69,47 73,25	51 - 90 55 - 86	1
Propylen	19	69,48	52 - 90	1
Butylen	19	68,73	52 - 90	1
Polyethylen (PE)	36	85,83	69 - 107	2
HDPE	10	80,98	69 - 102	
LDPE	22	88,55	73 - 107	
LLDPE	4	82,98	73 - 104	
Polypropylen	?	80,03	61 - 104	2
Benzen	5	73,67	53 - 84	3
Styren	7	98,39	80 - 108	3
Polystyren (PS)	20	102,16	83 - 114	3
GPPS ⁶	10	101,38	85 - 114	
HIPS ⁶	7	105,29	100 - 110	
EPS ⁶	3	96,22	83 - 100	
Chlor	14	21,5	10,3 - 25,6	4
Natriumhydroxid		20,7	7,2 - 23,7	
Hydrogen		17,2	6,9 - 22	
Vinylchlorid	10	56,8	43 - 65	4
PVC (gennemsnit)	14	66,8	48 - 89	4
Suspensions-PVC	8	64,9	48 - 75	
Emulsions-PVC	5	74,9	48 - 89	
Bulkpolym.-PVC	1	70,1	-	

1. Boustead (1993a).
2. Boustead (1993b).
3. Boustead (1993c).
4. Boustead (1994).
5. Antal anlæg er ikke givet: anlæggene forsynes med råstoffer via pipeline fra Nordsøen.
6. GPPS: "general purpose polystyrene", HIPS: "high impact polystyrene", EPS: ekspanderbart PS.

4.3.2 Sammenfattende vurdering

På tidspunktet for litteraturindsamlingen forelå der kun de europæiske plastproducenters data, hvorfor det ikke er muligt at sige noget generelt om industrielle data. Boustead (1992) bemærker, at de angivne intervaller ikke skal tages til udtryk for usikkerheder, men derimod som udtryk for varierende teknologi.

4.4 Generelle statistiske metoder i relation til livscyklusvurdering

4.4.1 Sammendrag

Publikationer af generelle statistiske metoder i relation til livscyklusvurderinger er ganske sparsomme. Der anvendes i stigende grad

simuleringer, blandt andet til livscyklusvurderinger. Beskrivelse af disse simuleringssmetoder vil kun blive givet i relation til livscyklusvurderinger, og ikke som en generel metode.

Simulering

Constantinou *et al.* (1992) beskriver anvendelse af simuleringssmodelen RiskPISCES (Constantinou & Seigneur, 1991) til risikovurdering i en helbredsundersøgelse i relation til kraftværker og emissioner herfra. Modellen benyttes til kombination og analyse af en række modeller, der beskriver transporten af forskellige kemiske komponenter i f.eks. luft, vand, jord, gennem fødekæden, etc. Usikkerhederne udtrykkes i denne model ved *Bayesianske* metoder, med angivelse af *a priori* sandsynligheder, og beregning af *posteriori* sandsynligheder. En række forskellige fordelinger er foreslægt for forskellige behandlede parametre (normalfordeling (f.eks. emissionsrate, depositions hastighed, jord/plante BCF), lognormal fordeling (f.eks. jorddybde, eksponeringens varighed), uniform fordeling (f.eks. tidsrum siden eksponeringens start, inhalations cancer potens faktor) og triangulær fordeling (f.eks. inhalationsrate)). Outputtet fra denne model er en deterministisk funktion af inputtet. Parametre optimeres ved mindste kvadraters metode.

Fremgangsmetoden er for hver enkel modelkomponent, Y at udvælge de uafhængige variabler blandt en række uafhængige variabler X_1, \dots, X_n , der forklarer den største del af variation for den afhængige variabel, Y . Når de vigtigste uafhængige variabler for hver enkel modelkomponent er udvalgt, opstilles en responsfunktion for hver modelkomponent, som funktion af de udvalgte uafhængige variabler. Denne responsfunktion estimerer den aktuelle model, som ofte er for tidskrævende at anvende, eller meget kompleks. Et eksempel på en sådan responsfunktion er følgende:

$$Y = f(X_1, \dots, X_k) \quad (3)$$

hvor $f(X_1, \dots, X_k)$ beskriver en funktion af variablerne X_1, \dots, X_k . Et eksempel på en simpel funktion kan f.eks. være en regressionsmodel.

Heresiter vælges en sandsynlighedsfordeling for hver af parametrene i modellerne.

De enkelte modeller kombineres nu i Monte Carlo simuleringer, hvor en sandsynlighedsfordeling estimeres for den afhængige variabel.

Metoden her beskrevet svarer til simuleringssprogrammet RISK (Palsade Corporation, 1991). RISK er nærmere beskrevet i afsnit 6.1.2.

Følsomhedsanalyse

Der er en lang række forskellige begreber for følsomhedsanalyse i den engelske litteratur, såsom f.eks. "sensitivity analysis" og "uncertainty analysis". I den danske litteratur anvendes udtryk som følsomhedsanalyse, fejlanalyse og usikkerhedsanalyse.

Analysemетодerne synes tilsvarende at være lidt forskellige. Nogle synes at foretage simuleringer, mens andre foretager beregning af størrelsen af usikkerheden. Nedenfor vil nogle af disse forskellige metoder blive beskrevet.

Rønning *et al.* (1993) beskriver usikkerhedsanalyse baseret på anvendelse af Gauss' usikkerhedsprincip (fejlophobning). Usikkerheden beregnes for en størrelse, X givet ved en funktion af en række variable

$$X = F(a, b, c, \dots) \quad (4)$$

hvor a, b, c ... er målte variabler. Usikkerheden beregnes ved skøn af usikkerheder på enkeltprocesser

$$\Delta x = \sqrt{\left(\frac{\partial x}{\partial a} \Delta a\right)^2 + \left(\frac{\partial x}{\partial b} \Delta b\right)^2 + \left(\frac{\partial x}{\partial c} \Delta c\right)^2 + \dots} \quad (5)$$

Heijungs (1995) foreslår en følsomhedsanalyse svarende til Rønning *et al.* (1993), men med en lille forskel. Heijungs (1995) udtrykker, som Rønning *et al.* (1993), en variabel X afhængig af en række uafhængige variabler, a, b, c (4).

Men Heijungs (1995) bestemmer usikkerheden ved

$$\Delta x = \left| \frac{\partial x}{\partial a} \right| \Delta a + \left| \frac{\partial x}{\partial b} \right| \Delta b + \left| \frac{\partial x}{\partial c} \right| \Delta c + \dots \quad (6)$$

Forskellen mellem Rønning *et al.* (1993) og Heijungs (1995) er, hvorvidt de enkelte led kvadreres før summation eller ej. Heijungs (1995) begrunder den direkte summation med, at man herved får et udtryk for den maximale fejl, hvorimod Rønning *et al.* (1993) får et udtryk for noget, der svarer til den forventede fejl. Herved er Heijungs' udtryk mere følsomt over for enkelte store fejl. Hvilket udtryk, der bør bruges, må til dels afhænge af formålet. Set fra en sælgers side er Rønning *et al.*'s udtryk at foretrække, da det giver mindste fejl. Som Heijungs nævner, må det være op til fremtidig diskussion inden for LCA, hvilket udtryk man bør benytte.

Heijungs *et al.* (1992) refererer, at udtrykket følsomhedsanalyse ofte benyttes til at bestemme betydningen af ændringer på slutresultatet. Heijungs (1994) beskriver 4 typer følsomhedsanalyse i forbindelse med livscyklusvurdering:

- Marginalanalyse
- Pålidelighedsanalyse
- Dominansanalyse
- Validitetsanalyse

Marginalanalyse og pålidelighedsanalyse er til dels én analyse. Ved en *marginalanalyse* vurderes betydningen af små ændringer på resultatet. Dette kan anvendes til at finde processer, hvor der i senere undersøgelser kan foretages forbedringer.

Pålidelighedsanalysen beskrives her som en analyse, der kan benyttes, når usikkerheden af procesdata kendes eller kan estimeres.

Dominansanalyse anvendes til vurdering af de mest dominerende led i en livscyklusvurdering, f.eks. hvilke(t) led der dominerer i

emissionen. Her er der ikke tale om en egentlig følsomhedsanalyse, da der ikke forekommer ændringer i data og følsomheden overfor variationer ikke indgår.

Validitetsanalysen anvendes til at vurdere, hvorledes valg og forudsætninger, foretaget i analysen influerer på resultatet. Dette har specielt betydning for vurdering af valg og forudsætninger, der er usikre.

Estimation af usikkerhed

Hoffmann *et al.* (1995) har beskrevet statistisk estimation af den samlede usikkerhed i 2 konstruerede livscyklusvurderinger. I livscyklusvurderingen beskrives den samlede proces i et flowdiagram. Hver enkelproces karakteriseres ved dens bidrag til den samlede proces med en middelværdi og en usikkerhed. Middelværdi og usikkerhed kan enten være estimerede på baggrund af forsøgsresultater eller skønnede størrelser på basis af f.eks. litteratur. Den samlede usikkerhed kan herefter estimeres ved antagelse om fordeling af de enkelte indgående variable i processen, samt koblingen mellem de enkelte led i processen. Hvis usikkerheden for processen eller enkelte dele i processen ikke kan estimeres eksakt, er der foretaget estimation ved anvendelse af Taylor's approximation. Således fås følgende udtryk for bestemmelse af variansen mellem 2 variabler A og B (Conradsen, 1984)

$$VAR(A+B) = VAR(A) + VAR(B) - 2\sqrt{VAR(A)VAR(B)}\rho \quad (7)$$

hvor ρ er korrelationen mellem A og B. Var (A) og Var (B) er variansen for henholdsvis A og B.

Hvis korrelationen mellem de 2 variable kan antages at være nul, kan udtrykket reduceres til

$$VAR(A+B) = VAR(A) + VAR(B), \quad \text{hvis } \rho = 0 \quad (8)$$

Endvidere fås tilsvarende for differencen mellem disse variable, med en korrelation på nul

$$VAR(A-B) = VAR(A) + VAR(B), \quad \text{hvis } \rho = 0 \quad (9)$$

Ved anvendelse af Taylor's approximation fås følgende udtryk for henholdsvis produktet og rationen mellem de 2 variable

$$VAR(AB) = \mu_B^2 VAR(A) + \mu_A^2 VAR(B) \quad (10)$$

$$\cdot \quad VAR\left(\frac{A}{B}\right) = \frac{1}{\mu_B^2} VAR(A) + \frac{\mu_A^2}{\mu_B^4} VAR(B) \quad (11)$$

hvor μ_A og μ_B er middelværdi for henholdsvis A og B.

En sammenligning mellem 2 eller flere processer kan foretages ved at bestemme et konfidensinterval for differencen mellem A og B

$$(\mu_A - \mu_B) \pm t(0,975; df) s \sqrt{\left(\frac{2}{n}\right)} \quad (12)$$

hvor μ_A og μ_B er middelværdi for henholdsvis A og B
 s er estimat af spredningen
 n er antal observationer for henholdsvis A og B
 df er antallet af frihedsgrader ved estimation af s
 $t(0,975; df)$ er værdi fra t-fordelingen med df frihedsgrader svarende til 95% fraktilen.

4.4.2 Sammenfattende vurdering

Litteraturen omkring statistik i relation til livscyklusvurderinger er yderst sparsom.

Metoder

Metoderne beskrevet i litteraturen kan deles op i to grupper: simulation (Constantinou *et al.* (1992)) og direkte estimation (Hoffmann *et al.* (1995)). Begge metoder er baseret på antagelser om de indgående variables fordelinger, middelværdi og standardafvigelse. På basis af en række inputvariabler bestemmes ved begge metoder middelværdi og standardafvigelse for de(n) indgående variable.

Vurdering

Til sammenligning og vurdering af et eller flere parallelle produkter benyttes ofte følsomhedsanalyser. Herved fås en vurdering af usikkerheden på outputvariablerne i relation til inputvariablerne. Rønning *et al.* (1993) og Heijungs (1995) har givet forskellige forslag til bestemmelse af usikkerheden på outputvariablerne ved brug af forskellige udgaver af Gauss' usikkerhedsprincip, hvor den første udgave giver et udtryk for den forventede fejl, mens den anden udgave giver et udtryk for den maximale fejl. Heijungs (1994) beskriver 4 typer følsomhedsanalyser, der f.eks. kan benyttes i forbindelse med en livscyklusvurdering (marginalanalyse, pålidelighedsanalyse, dominansanalyse og validitetsanalyse).

4.5 Opfattelse og betydning af usikkerheder

4.5.1 Sammendrag

"Spredning" er et veldefineret begreb i den statistiske terminologi, der angiver, hvilken usikkerhed (fejl) en given størrelse er behæftet med. Her kan henvises til generelle lærebøger i statistik samt ISO standard med statistisk terminologi (ISO 3534-1:1993; ISO (1993)).

Usikkerhed anvendes ligestillet med spredning og standardafvigelse.

Det har ikke været muligt at finde nogle egentlige kilder, der beskriver forskellige opfattelser af begrebet "usikkerhed" blandt faggrupper uden for det matematiske område. Ved almindelig forespørgsel i en række forskellige fagområder har det dog vist sig, at der er mange forskellige opfattelser af dette begreb.

Her gengives enkelte udsagn om usikkerheder:

1. "Usikkerhed betyder, at variablen med 95% sikkerhed ligger inden for intervallet $x - y$ "
2. "Der er en usikkerhed på 10%"
3. "Der er en usikkerhed på 2 MJ"

Det første udsagn udtrykker et konfidensinterval, hvorimod det andet udsagn formodes at udtrykke en variationskoefficient. Det tredje udsagn formodes at udtrykke en absolut størrelse som en spredning.

Der forekommer et eksempel på software, som indeholder muligheder for behandling af usikkerheder i forbindelse med livscyklusvurdering (KCL-ECO) (The Finnish Pulp and Paper Research Institute (KCL), 1995). I nævnte program svarer den usikkerhed, der kan tilknyttes de enkelte inputdata, til tre gange standardafvigelsen (3σ).

4.5.2 Sammenfattende vurdering

Det er påfaldende, at der tilsvneladende ikke er nogen, der er gået dybere ind i undersøgelse af den generelle opfattelse af begrebet usikkerhed inden for forskellige områder.

4.6 Sammenfatning af erfaringer fra videnindsamlingen

Litteratursøgningerne resulterede kun i få referencer vedrørende statistiske metoder. Heraf fremkom der relativt mange referencer vedrørende usikkerheder i livscyklusvurderinger. I de beskrevne metoder er usikkerhederne forsøgt karakteriseret efter deres oprindelse (kilde, lande, alder, ...). Der blev kun fundet oplysninger om usikkerheder på industri- og miljødata fra plastindustrien, som har publiceret europæiske gennemsnitsdata for produktion af en række plastmaterialer. De fundne referencer vedrørende generel anvendelse af statistisk analyse omhandler simuleringer eller mere direkte estimation samt forskellige former for følsomhedsanalyser. Opsattelsen af usikkerheder i de forskellige referencer er meget varierende og pegede ikke på én fælles standard.

5. Datakvalitet

5.1 Introduktion til beskrivelse af datakvalitet

5.1.1 Data og datakvalitet

Data

Data, der anvendes i livscyklus-studier, er *miljødata* vedrørende de undersøgte processer, *systemdata* vedrørende flowet af råstoffer, energi og produkter gennem de undersøgte processer, samt *ydelses-data* vedrørende den funktionelle produktenhed, der anvendes ved sammenligning af forskellige produktionssystemer. Når der foretages en komplet livscyklusvurdering, tilføjes yderligere data til at *karakterisere* og *vurdere* miljøpåvirkninger. Data skal altså forstås som enhver oplysning, der indgår i en livscyklusvurdering (Hoffmann *et al.*, 1995).

Datakvalitet

Datakvalitet er de specifikke egenskaber ved data, udtrykt gennem meta-data (information om data), f.eks. oplysninger om data-usikkerhed (spredning og fordelingsmønster), data-pålidelighed (afhængig af de metoder, der anvendes ved måling, beregning, skøn og kvalitetsbedømmelse af data), graden af fuldstændighed (antal dataindsamlingssteder og perioder og deres repræsentativitet for hele populationen), data-alder (år for oprindelig måling), det geografiske område, for hvilket data er repræsentative, samt den proceseknologi eller det teknologiske niveau, for hvilket data er repræsentative.

Pålideligheden af resultatet af en livscyklusvurdering, og dermed anvendeligheden af dette resultat, afhænger af kvaliteten af de oprindelige data, der ligger til grund for vurderingen. Beskrivelse af datakvalitet må derfor ses som en integreret del af livscyklusvurderingen.

Beskrivelse af datakvalitet

En formaliseret beskrivelse af datakvalitet begynder med en fastlæggelse af en målsætning for datakvalitet og en dataindsamlingsstrategi, som led i målformuleringen og afgrænsningen af den enkelte livscyklusvurdering. Under dataindsamlingen dokumenteres datakvalitet for hvert sæt data. Kvaliteten af de enkelte datasæt kan derefter relateres til datakvalitetsmålene ved hjælp af et antal datakvalitets-indikatorer, der angiver datakvaliteten i forhold til den måde, data anvendes i undersøgelsen. Endelig kan datakvalitets-indikatorerne fortolkes som en yderligere usikkerhed på de enkelte data. Sådanne supplerende usikkerheder kan inddrages, når man beregner usikkerheden på det samlede resultat af livscyklusvurderingen. De kan således anvendes til at give et mere kvalificeret udtryk for dette resultat.

I denne rapport anvendes den beskrevne fremgangsmåde på udvalgte dele af et livscyklus-studie af rugbrød fra Weidema *et al.* (1995).

Datakvalitets-index

Vi indfører brugen af et datakvalitets-index (inspireret af Funtowicz & Ravetz 1990), der giver semi-kvantitative oplysninger vedrørende:

- Pålideligheden af data (uafhængig af datas anvendelse i den konkrete undersøgelse), herunder en vurdering af indsamlingsmetoder og verifikationsprocedurer.
- Fuldstændigheden af data (uafhængig af datas anvendelse i den konkrete undersøgelse), herunder den statistiske repræsentativitet, antal indsamlingspunkter og tidsrum for dataindsamling.
- Datas tidsmæssige, geografiske og teknologiske korrelation (i forhold til den konkrete undersøgelses målsætning for datakvalitet).

Datakvalitets-indikatorer

Disse indikatorer for datakvalitet kan derefter anvendes:

- til at revidere dataindsamlings-strategien med henblik på at forbedre kvaliteten af de indsamlede data;
- i kombination med skøn for usikkerhed, til at give en bedre vurdering af resultatets pålidelighed.

5.1.2 Datakvalitetsmålsætning

Datakvalitetsmålsætningen angiver i generelle vendinger de ønskede specifikationer for de data, der ønskes anvendt i undersøgelsen. En datakvalitetsmålsætning kan udtrykke:

- den relative eller absolute betydning af forskellige aspekter vedr. datakvalitet (f.eks. verifikation, datas alder),
- det ønskede geografiske og teknologiske niveau,
- betydningen af at anvende data af sammenlignelig kvalitet til de forskellige produkt-alternativer, der sammenlignes,
- håndtering af manglende data og manglende repræsentativitet (f.eks. om der skal bruges skøn for værst tænkelige udfald eller skøn for forventede gennemsnit),
- håndtering af manglende oplysninger om usikkerhed på data.

Krav til datakvalitet

En datakvalitetsmålsætning kan have vidt forskellige krav til datakvalitet for de forskellige dele af livscyklus (de dele, der har stor betydning for resultatet, skal dækkes ind med data af bedre kvalitet end de dele, der har mindre betydning). Det vil sige, at datakvalitetsmålsætningen ikke opstiller detaljerede krav til de enkelte datasæt, men er en generel angivelse af de fremgangsmåder og prioriteringer, der skal anvendes ved dataindsamling og bedømmelse af datakvalitet.

Formål og anvendelse

Datakvalitetsmålsætningen for en specifik undersøgelse afhænger af undersøgelsens formål og anvendelsesområde f.eks.:

- en virksomheds-specifik undersøgelse til brug for langsigtet strategi vil forholde sig til andre geografiske og teknologiske niveauer end en offentlig undersøgelse, der skal sammenligne eksisterende produkter på markedet
- en virksomheds-specifik undersøgelse vil lægge større vægt på at benytte de nyeste data end på at benytte data, der er verificerede, mens der måske prioriteres omvendt ved en offentlig undersøgelse
- anvendelse af sammenlignelig datakvalitet vil være af størst betydning for den sammenlignende undersøgelse

- i tilfælde af manglende data vil den sammenlignende undersøgelse som regel støtte sig til skøn over værst tænkelige tilfælde, mens den strategiske undersøgelse evt. blot vil anføre, at de pågældende data mangler

Eksempel på mål

I eksemplet med rugbrød (se de følgende afsnit i kapitel 5) blev data-kvalitetsmålsætningen angivet som følger:

- nye data (helst for høståret 1992) har højere prioritet end andre datakvalitets-aspekter;
- virksomhedsspecifikke data for det aktuelle teknologiske niveau foretrækkes;
- skøn for værst tænkelige tilfælde anvendes, når data er mangelfulde m.h.t. tilgængelighed eller repræsentativitet;
- når oplysninger vedr. usikkerhed mangler, skal der anvendes standardskøn over usikkerheden.

5.1.3 Indsamling af data

Dokumentation

Den egentlige dokumentation af datakvalitet fremkommer som en del af dataindsamlingsprocessen:

- indsamlingsmetode (målemetode for måledata, beregningsmetode for gennemsnit eller øvrige beregnede data, angivne forudsætninger for skønnede data)
- de metoder, der anvendes til verifikation af data
- antal dataindsamlingspunkter og -perioder, samt deres repræsentativitet i forhold til den samlede population
- data-alder, udtrykt som år for den oprindelige måling
- det geografiske område, som de pågældende data er repræsentative for
- den proceseknologi eller det teknologiske niveau, for hvilke data er repræsentative

Nogle aspekter af datakvaliteten (f.eks. data-alder samt geografiske og teknologiske specifikationer) er typisk relateret til den generelle proces (systemdata), snarere end til de enkelte miljørelationer og kan deraf rapporteres i den overordnede procesbeskrivelse. Når datakvaliteten ikke kendes, bør den angives som ukendt.

Eksempel på rapportering

I eksemplet med rugbrød (se de følgende afsnit i kapitel 5) findes et datasæt for den mængde pesticider, der anvendes til de danske afgrøder. Den blev opgivet til at være 1,9 kg aktivt stof/ha, beregnet ud fra salgsstatistikkerne fra 1992, divideret med det samlede landbrugsareal i 1992 (begge tal fra den danske landbrugsstatistik, 1993). Det samlede landbrugsareal er bestemt ved at sammentælle alle ejendomme over 5 ha (100% repræsentativitet, idet ejendomme på under 5 ha udgør mindre end 0,5% af det samlede antal), og den solgte pesticidmængde indberettes af alle danske producenter og importører (142 virksomheder = 100% repræsentativitet) og omfatter også de mængder, der anvendes i private haver; den tager dog ikke højde for lagerforskydning i handelen over år. Data fra statistiske kontorer anses i almindelighed som tilstrækkeligt verificeret.

Energiforbruget til pesticidsfremstilling blev opgivet som 224 MJ/kg aktivt stof, baseret på en teoretisk beregning foretaget af Green & McCulloch (1976) og Green (1987) delvis baseret på oplysninger i patentansøgninger, omfattende værdier for 39 pesticider med et samlet energiforbrug spændende fra 58 til 580 MJ/kg, i gennemsnit 240 MJ/kg, hvoraf 77 MJ er kul til elektricitet ved en effektiv udnyttelse på 31%. Værdien på 224 MJ/kg fremkom ved at justere til en nutidig effektiv udnyttelse på 42% og tillægge 5 MJ brændstof/kg aktivt stof (Green 1976) for at tage højde for fyld- og bærestoffer. De pågældende data er ikke verificeret.

5.1.4 Indikatorer for datakvalitet

Kvaliteten af de individuelle data kan derefter relateres til målene for datakvalitet ved hjælp af en række datakvalitets-indikatorer, der angiver datakvaliteten i forhold til den måde, data anvendes i undersøgelsen. Indikatorer for datakvalitet kan anvendes til at bedømme overensstemmelsen mellem de angivne mål for datakvalitet og de indsamlede datas faktiske kvalitet.

Datakvalitetsindikatorer er semi-kvantitative værdier, der knyttes til et sæt data, og som repræsenterer de pågældende datas kvalitet. Vi har kunnet konstatere, at de følgende indikatorer for datakvalitet er nødvendige og tilstrækkelige til at beskrive de aspekter ved datakvalitet, der påvirker resultatets pålidelighed (se også tabel 5.1.1.).

Pålidelighed

"Pålideligheds-indikatoren" er relateret til de kilder, indsamlingsmetoder og verifikationsprocedurer, der anvendes for at fremskaffe de pågældende data. Indikatoren er uafhængig af datakvalitets-målsætningen: et krav, der fastlægges i målsætningen ændrer ikke datas pålidelighed.

Fuldstændighed

"Fuldstændigheds-indikatoren" er relateret til datas statistiske egenskaber: Hvor repræsentativ prøven er, om prøven indeholder et tilstrækkeligt antal data, og om perioden er tilstrækkelig til at udjævne normale udsving. Ligesom "pålideligheds-indikatoren" er "fuldstændigheds-indikatoren" uafhængig af datakvalitetsmålsætningen.

De følgende tre indikatorer vedrører alle sammenhængen mellem data og de datakvalitets-mål, der vedrører produktionsforholdene i bred forstand:

Tidsmæssig korrelation

"Tids-indikatoren" udtrykker den tidsmæssige korrelation mellem undersøgelsesår (som angivet i datakvalitetsmålsætningen) og indsamlingsår for de fundne data. Da den teknologiske udvikling går meget stærkt i visse industrier, kan en forskel på 10 år mellem undersøgelsesår og dataindsamlingsår betyde, at udledninger og produktionens effektivitet når at ændre sig totalt. Det betyder, at "tids-indikatoren" er nært knyttet til datakvalitetsmålsætningen.

Geografisk korrelation

Den "geografiske indikator" udtrykker den geografiske korrelation mellem undersøgelsesområdet (som opgivet i datakvalitets-målsætningen) og de anvendte data. Produktionsmetoder og -betingelser kan være meget forskellige i Norge (f.eks. landbrug med små brug), USA

(stordrift, moderne teknologi) og landene i det østlige Europa (ældre teknologi).

Teknologisk korrelation

Den "teknologiske indikator" vedrører alle andre aspekter end de tidsmæssige og geografiske overvejelser. Selv om data kan have den ønskede alder og være repræsentative for det ønskede geografiske område, kan de muligvis ikke fremskafles for de specifikke virksomheder, processer eller materialer, som undersøgelsen vedrører. Derfor kan det være påkrævet at anvende data fra beslægtede processer eller materialer, hvilket i visse tilfælde kan være at foretrække frem for ældre data eller data fra et andet geografisk område.

Indbyrdes uafhængighed

Det er vigtigt at de forskellige aspekter af datakvaliteten, der karakteriseres af de fem indikatorer, betragtes som indbyrdes uafhængige. Det er oplagt, at data fra et "forkert" område kan indebære, at der anvendes en anden teknologi. Ikke desto mindre er dette aspekt beskrevet under "geografisk korrelation". Den teknologiske korrelation vedrører kun de yderligere forskelle, der optræder, selv om data er af *samme* alder og fra *samme* geografiske område.

På samme måde kan fuldstændigheds-indikatoren angive fuldstændig repræsentativitet, selv når de tre korrelations-indikatorer angiver meget dårlig korrelation. Det skyldes, at repræsentativiteten ikke henviser til den undersøgelse, hvori de pågældende data anvendes, men kun til data selv. Et sæt data kan være fuldstændig repræsentativt for situationen i Storbritannien i 1976, men alligevel have en meget dårlig korrelation, hvis undersøgelsen vedrører den franske industri i 1995. Omvendt kan det forekomme, at et sæt data fra den undersøgte virksomhed, der er up-to-date og passer perfekt, ikke er fuldstændigt.

Tabel 5.1.1

Datakvalitetsmatrix med 5 datakvalitetsindikatorer (let revideret efter Weidema 1994).

Indikator score	1	2	3	4	5
Pålitelighed	Verificerede data baseret på målinger	Verificerede data delvis baseret på antagelser eller ikke-verificerede data baseret på målinger	Ikke-verificerede data delvis baseret på antagelser	Kvalificeret skøn (f.eks. af sagkyndig ekspert fra industrien)	Ikke-kvalificeret skøn
Fuldstændighed	Repræsentative data fra et tilstrækkeligt antal indsamlingssteder over et tidsrum der er tilstrækkeligt til at udjævne normale udsving	Repræsentative data fra et mindre antal antal indsamlingssteder, men for tilstrækkelige perioder	Repræsentative data fra passende antal indsamlingssteder, men for kortere tidsrum	Repræsentative data, men fra mindre antal steder og kortere tidsrum eller ufuldstændige data fra et tilstrækkeligt antal indsamlingssteder og tidsrum	Repræsentativitet ukendt eller ufuldstændige data fra et mindre antal indsamlingssteder og/eller for kortere tidsrum
Tidsmæssig korrelation	Mindre end 3 års forskel til undersøgelsesår	Mindre end 6 års forskel	Mindre end 10 års forskel	Mindre end 15 års forskel	Dataalder ukendt eller mere end 15 års forskel
Geografisk korrelation	Data fra undersøges-område	Gennemsnitsdata fra større område, der også omfatter undersøges-området	Data fra område med lignende produktions-betingelser	Data fra område med beslægtede produktions-betingelser	Data fra ukendt område eller område med meget anderledes produktionsbetingelser
Teknologisk korrelation	Data for virksomheder, processer og materialer omfattet af undersøgelsen	Data for processer og materialer omfattet af undersøgelsen, men fra andre virksomheder	Data for processer og materialer omfattet af undersøgelsen, men fra anden teknologi	Data for beslægtede processer eller materialer, men fra samme teknologi	Data vedr. beslægtede processer eller materialer, men fra anden teknologi

Formålet med datakvalitetsmatricen i tabel 5.1.1. er at tilvejebringe et redskab til styring af datakvalitet, der gør det let at overske datakvaliteten for dermed at kunne pege på muligheder for forbedringer i datakvalitet samt efterspore kilder til usikkerheder.

Scorerne i datakvalitetsmatricen er "semi-kvantitative". De tjener udelukkende som identifikationsværdier og må ikke betragtes som noget, der repræsenterer en bestemt "mængde" datakvalitet. Derfor bør værdierne ikke opsummeres.

Eksempler på bestemmelse af datakvalitetsindikatorer

I eksemplet med rugbrød kan pesticidsforbruget for danske afgørder opgives som 2 kg aktivt stof pr. hektar. Når man undersøger baggrunden for disse forbrugsdata, kan dets datakvalitetsindeks bestemmes som (2,1,1,2,4). Forklaringen er vist i tabel 5.1.2.

Tabel 5.1.2

Datakvalitetsindikator for mængde anvendt pesticid.

Datakvalitetsindikator	Datakvalitet	Score
Pålidelighed	Datakilden er Danmarks Statistik, hvis data er baseret på indberetninger fra importører og producenter. Brugen i private haver og lagerforsyning i de pågældende år er ikke målt, men antages at være af mindre betydning.	2
Fuldstændighed	Data bestemmes for Danmark i perioden 1988 til 1992. Dette er tilstrækligt til at dække udsvingning i de pågældende år (jordbrugetsforholdene skifter med vejret, og dermed fra år til år)	1
Tidsmæssig korrelation	Data dækker undersøgelsesåret	1
Geografisk korrelation	Data er et gennemsnit for Danmark, ikke for de danske landbrug, der faktisk leverer ingredienserne til rugbrødet.	2
Teknologisk korrelation	Data er et gennemsnit for alle danske afgørder, ikke for de faktiske afgørder, der anvendes som ingredienser i rugbrødet.	4

På tilsvarende måde var energiforbruget til fremstilling af pesticider, 240 MJ/kg aktivt stof. Vi kan tilføje det beskrivende indeks (3,1,5,3,4), der bestemmes som vist i tabel 5.1.3.

Tabel 5.1.3

Datakvalitetsindikator for energiforbrug til fremstilling af pesticid.

Datakvalitetsindikator	Datakvalitet	Score
Pålidelighed	Datakilden anvender computerberegninger, der delvis er baseret på procesbeskrivelser i patentansøgninger, og ikke på målte data.	3
Fuldstændighed	Data er som sådan repræsentative for de 39 pesticider omtalt i kilden.	1
Tidsmæssig korrelation	De oprindelige data i computerprogrammet er fra 1976 og er baseret på ældre data (datakvalitets-målsætningen var 1992).	5
Geografisk korrelation	Generelle data anvendes for de specifikke pesticider, der anvendes i Danmark. Produktionsbetingelserne for pesticider i forskellige geografiske områder bedømmes til at være forholdsvis ens med hensyn til energiforbrug.	3
Teknologisk korrelation	Et gennemsnit for de 39 pesticider i kilden anvendes for alle de pesticider, der er medtaget i undersøgelsen.	4

Ved at opstille datakvalitetsindikatorer for alle de individuelle data i en undersøgelse kan man få et bedre indtryk af de typiske problemer, der er forbundet med datakvaliteten i en given undersøgelse. Det er særdeles nyttigt, når man vil forbedre dataindsamlingsstrategien i løbet af et livscyklus-studie.

Når data kombineres til en værdi for den samlede undersøgelse (f.eks. energiforbruget til 1 kg rugbrød), vil datakvalitetsindeks for den samlede værdi dog typisk være (5,5,5,5,5), idet den aldrig kan være bedre end datakvalitetsindekset for de data, der indgår i den samlede værdi, og det vil sige, at den kun afspejler den dårligste score for hver datakvalitets-indikator. Hvis der kun bliver opgivet den slags samlede værdier, vil det ikke være særlig oplysende, idet det kun siger noget om, at der forekommer et problem med datakvaliteten et eller andet sted i undersøgelsen, men ikke noget om hvor og hvor alvorligt problemet er. Hvis man således udelukkende anvender datakvalitetsindekser, opnås ikke tilstrækkelig oplysning om kvaliteten af livscyklusvurderingens *resultat*. For at rette op på denne begrænsning må man kombinere indekserne med oplysninger om usikkerhed for de enkelte data.

5.1.5 Modifikation af kvalitetsindikatorer

Vi skelner i det følgende mellem to typer datausikkerhed:

- den *oprindelige usikkerhed*, der er forbundet med alle indsamlede data (typisk målefejl og normale udsving i den målte variabel), og
- en *supplerende usikkerhed*, der skyldes, at data ikke er af optimal kvalitet, hvilket afspejles i et datakvalitetsindeks forskelligt fra (1,1,1,1,1).

Ofte kender man ikke den oprindelige usikkerhed. Hvis der foreligger flere målte værdier, har man mulighed for at beregne datas oprindelige usikkerhed, men i reglen findes der kun én værdi.

Datakvalitetsindikatorerne beskriver kun data, som de fremstår med deres oprindelige usikkerheder. Hvis man derudover tager højde for et bestemt datakvalitetsaspekt, når man beskriver datas usikkerhed, bør kvalitetsindekset for det pågældende data modificeres tilsvarende. For at vise, at et nyt indeks ikke henviser til de oprindelige data, men til de *modificerede* data, kan anvendes et "m" i stedet for den normale datakvalitetsindikator.

Modifikation af indeks

I ovennævnte eksempel blev den anvendte pesticidmængde til danske afgrøder opgivet som 2 kg aktivt stof/ha, baseret på salgsstatistikker. Den basale usikkerhed i dette tal kan bestemmes ud fra den årlige variation i salgsstatistikkerne. Ved at bruge statistikkerne fra 1988 til 1992 får man en variationskoefficient på 5,5%. Den tillægsusikkerhed, der hidrører fra datakvalitetsindekset (2,1,1,2,4), kan vurderes som vist i tabel 5.1.4.

Tabel 5.1.4

Datakvalitets-indikator for anvendt mængde pesticid med supplerende usikkerhed som følge af dårlig score på én eller flere datakvalitetsindikatorer.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed ¹
Oprindelig usikkerhed	Variation mellem år		CV: 5,5%
Pålidelighed	Skøn for værst tænkelige tilfælde af forbrug i private haver og lagerforskydning over år.	2	CV: 2,5%, estimat
Fuldstændighed	-	1	-
Tidsmæssig korrelation	-	1	-
Geografisk korrelation	Estimeret variation mellem bedrifter for samme afgrøde	2	CV: 17%, estimat
Teknologisk korrelation	Estimeret variation mellem afgrøder på samme bedrift.	4	CV: 33%, estimat
Samlet usikkerhed			CV: 38%

CV: Variationskoefficient.

Den samlede usikkerhed beregnes som en variationskoefficient (kvadratroden uddrages af summen af kvadratet på de enkelte koefficenter). Det vil sige, at de 2 kg/ha (CV: 5,5%)(2,1,1,2,4) nu kan skrives som 2 kg/ha (CV: 38%)(m,1,1,m,m).

Tilsvarende blev energiforbruget til fremstilling af pesticider opgivet til 240 MJ/kg aktivt stof, med en variationskoefficient på 58,7% (3,1,5,3,4). Den supplerende usikkerhed kan vurderes som vist i tabel 5.1.5.

Tabel 5.1.5

Datakvalitets-indikator for pesticidfremstilling med supplerende usikkerhed som følge af dårlig score på én eller flere datakvalitetsindikatorer.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed ¹
Oprindelig usikkerhed	Beregnet variation i de 39 pesticider anført i kilden.		CV: 58,7%
Pålidelighed	Estimeret forskel mellem patentansøgninger og faktisk produktion.	3	CV: 17%, estimat
Fuldstændighed	Som sådan er data repræsentative for de 39 pesticider omtalt i kilden.	1	-
Tidsmæssig korrelation	Estimeret stigning i energi-effektivitet	5	reducér middelværdi med 40%
Geografisk korrelation	Estimeret variation mellem forskellige produktions-anlæg.	3	CV: 10%, estimat
Teknologisk korrelation	Estimeret øget energiforbrug til moderne komplekse pesticider, f.eks. pyrethroider, sammenlignet med ældre, mere simple typer.	4	forhøj middelværdi med 60%
Samlet usikkerhed			CV: 62%; forøg middelværdi med 20%

1. Variationskoefficenter, med mindre andet er anført.

Energiforbruget til fremstilling af pesticider bliver herved 288 MJ (CV: 62%)(m,1,m,m,m).

Som det fremgår af det foregående eksempel, kan en lav score på datakvalitetsindikatoren både forhøje usikkerheden og ændre middelværdien. Usikkerhederne for de enkelte datakvalitetsindikatorer er additive, idet indikatorerne er bestemt som indbyrdes uafhængige.

Supplerende usikkerheder

I nogle tilfælde er det muligt at beregne den supplerende usikkerhed. I andre tilfælde er man nødt til at estimere usikkerheden. I de foregående eksempler er der anført specifikke estimer på grundlag af teknisk viden om de pågældende processer. En anden mulighed kunne være at anvende standardskøn for usikkerhed, der kan opstilles ud fra andre, lignende datasæt, for hvilke usikkerhederne kendes.

Tabel 5.1.6 viser en matrix over usikkerheder, hvor de typiske supplerende usikkerheder fra case-studiet af rugbrød (Weidema et al., 1995, kap. 5.2.25) er indført. De fleste usikkerheder er skønnede. Hvis et større antal beregnede værdier for disse supplerende usikkerheder var til rådighed, ville det være muligt at anvende en sådan matrix til at tildele standard-usikkerheder til hver score for hver datakvalitets-indikator. Dette bør være genstand for yderligere undersøgelser. Det må forventes, at det vil være nødvendigt at opbygge forskellige "standard-usikkerheds-matricer" for forskellige emner (til forskellige typer data, forskellige miljørelationer og forskellige typer industri).

Tabel 5.1.6

Påbegyndt matrix til akkumulering af typiske supplerende usikkerheder (som varianskoefficienter, med mindre andet er oplyst) for de enkelte datakvalitets-indikatorer. Værdierne i matricen er hovedsagelig baseret på estimer. De henviser kun til materialestrømme og energiforbrug og er kun anført som eksempel. Værdierne bør ikke anvendes som standardværdier i andre studier, før der er indhøstet flere erfaringer.

Indikator score	1	2	3	4	5
Pålidelighed	0%	3 - 10%	17 - 25%	25%	97%
Fuldstændighed	0%	0 - 10%	?	?	25%
Tidsmæssig korrelation		ved energiforbrug: reducér middelværdi med:			
	0%	10%	20%	30%	40%
Geografisk korrelation (lave værdier i industrien, høje værdier i landbruget)	0%	5 - 17%	10-25%	50%	50%
Teknologisk korrelation	0%	16%	25%	33%	50%

Det forventes, at de erfaringer, der opnås ved at udarbejde sådanne "standard-usikkerheds-matricer" også vil føre til en forbedret definition og afgrænsning af de enkelte scorer i den generelle datakvalitetsmatrix (tabel 5.1.1.).

5.2 Case 1: Energiforbrug til landbrugsafgrøder

5.2.1 Beskrivelse af casen

Casen tager udgangspunkt i et konkret produkt - et skiveskåret rugbrød fra en specifik dansk fabrik. Data er taget fra en publiceret livscyklus-screening af dette produkt (Weidema et al., 1995).

I et rugbrød indgår en række landbrugsafgrøder. Først og fremmest rug og hvede, hvoraf en del anvendes som hele kerner og en del i

form af mel. Dertil kommer typisk en række eksotiske frø (her hørfrø, sesamfrø og solsikkekerner), gær (fremstillet af melasse fra sukkerroer) og slipmiddel (palmeolie fremstillet af palmefrugt). Det samlede energiforbrug til disse landbrugsafgrøder omfatter energiforbruget til landbrugets indsatsstoffer (kemikalier og fødning), og den energi der bruges i selve landbruget (til markarbejde og tørring af afgrøderne). Disse energiforbrug, der typisk opgøres i MJ primær energi pr. hektar ($ha = 10.000 m^2$), skal - via høstudbyttet pr. ha - relateres til de mængder af de enkelte afgrøder, der indgår i rugbrødsproduktionen.

Rugbrøds-eksemplet er valgt til at belyse typiske usikkerheder på data fra mange producenter med stor geografisk spredning. Data er ofte gennemsnit fra statistiske opslagsværker og dækker ikke præcis det undersøgte produkt. Dersør bliver der ofte tale om meget store usikkerheder som følge af dårlig datakvalitet.

I de oprindeligt publicerede data (Weidema *et al.*, 1995) er høstudbyttene og dermed arealsforbruget kun opgjort for 1992. Dette medfører for disse data en lav score (3) for datakvalitetsindikatoren: fuldstændighed, idet data fra et enkelt år ikke kan siges at dække de normale variationer i udbyttet, der er stærkt afhængig af de klimatiske forhold i de enkelte år. I nærværende beskrivelse af casen er det valgt at benytte data for flere år for dermed at få et udtryk for den faktiske usikkerhed som følge af disse års-variationer. Dermed ændres scoren for fuldstændighed fra 3 til 1 (data er nu repræsentative for de normale variationer), mens usikkerheden angives direkte i tilknytning til data.

I tabel 5.2.1 er givet en oversigt over data, der indgår i casen. Datakvaliteten ses i tabel 5.2.25 i afsnit 5.2.3. En detaljeret beskrivelse af datakvaliteten findes i det mellemliggende afsnit 5.2.2.

Tabel 5.2.1
Oversigt over data.

Data	Kommentarer
Afgrøder til 1 kg rugbrød.	Opskrift fraregnet tab i møllerier og sukkerfabrik.
Høstudbytte pr. ha af de forskellige afgrøder.	Gennemsnit for flere år for oprindelseslandet undtagen for palmefrugt, hvor der er brugt specifikke data fra en Malaysisk plantage i 1990.
Forbrug af pesticider pr. ha.	Gennemsnit for totalt forbrug af pesticider i Danmark divideret med totalt dyrket areal. Pesticidforbruget for de enkelte afgrøder er ikke kendt.
Energiforbrug til produktion af pesticider.	Gennemsnit af teoretiske beregninger fra 1976 for 39 pesticider delvist baseret på oplysninger i patentanmeldelser.
Forbrug af N, P og K kunstgødning pr. ha pr. dansk afgrøde.	Det totale forbrug af N, P og K kunstgødning i Danmark allokeret på samtlige danske afgrøder i forhold til den anbefalede mængde af N, P og K i gødning pr. afgrøde.
Forbrug af N, P og K kunstgødning pr. ha pr. udenlands afgrøde.	Gennemsnit for totalt forbrug af N, P og K kunstgødning for det samlede dyrkede areal i oprindelseslandet.
Energiforbrug til produktion af N kunstgødning.	For danske afgrøder er anvendt data for ammoniumnitrat og sammensatte gødninger fra Norsk Hydro fra 1992 opgivet som intervaller, der antages at dække forskelle i aktuel teknologi mellem forskellige producenter. For udenlandske afgrøder anvendes ovennævnte data og tilsvarende data for urea i et forhold som varierer fra land til land.
Energiforbrug til produktion af P kunstgødning.	Data fra Norsk Hydro fra 1992 opgivet som intervaller, der antages at dække forskelle i aktuel teknologi.
Energiforbrug til produktion af K kunstgødning.	Data for Norsk Hydro fra slutningen af 1980'erne.
Forbrug af kalk pr. ha pr. afgrøde.	Gennemsnit af det totale danske forbrug 1983-92 allokeret på samtlige kalkkrævende afgrøder.
Energiforbrug til fremstilling af kalk.	Estimat baseret på review af flere kilder.
Forbrug af husdyrgødning pr. ha pr. afgrøde.	Den nødvendige mængde husdyrgødning til at dække forskellen mellem forbruget af N i kunstgødning og de anbefalede mængder N for samtlige danske afgrøder ved en udnyttelsesgrad på 40% af N i husdyrgødning. Der anvendes ikke husdyrgødning til udenlandske afgrøder.
Energiforbrug til fremstilling af husdyrgødning.	Gennemsnitlige kapacitetsomkostninger til ei pr. svin divideret med elprisen og mængden af husdyrgødning pr. svin.
Energiforbrug pr. ha til markarbejde for danske afgrøder.	Brændstofforbrug for udvalgte maskiner målt over flere år ganget med det målte timeforbrug til forskellige markarbejder. For de maskintyper hvor brændstofforbruget ikke er målt i praksis er det antaget at svare til forbruget ved 80% af maksimal belastning.
Energiforbrug til markarbejde pr. ha for udenlandske afgrøder.	Målte eller beregnede data for forskellige afgrøder i 18 lande.
Energiforbrug til tørring af korn-afgrøder.	Kvalificeret estimat baseret på gennemsnitligt vandindhold i korn.

5.2.2 Beskrivelse af datakvalitet

Afgrøder til 1 kg rugbrød

Forbruget af råvarer pr. kg solgt rugbrød er beregnet ud fra fabrikvens opskrift justeret for tab i pakkemaskine og intern recirkulering af brødkrummer. Omregningen til afgrøder (se tabel 5.2.2.) sker ved at tillægge tabet i møllerier, gærfabrik og sukkerfabrik ud fra de oplyste formalingsprocenter (85% for rugmel og 75% for hvedemel) og udbytter (0,22 liter palmeolie pr. kg palmefrugt, 0,833 kg gær pr. kg

melasse og 0,035 kg melasse pr. kg sukkerroe). Øvrige tab rapporteres som forsvindende (< 0,1%).

Usikkerheden på tallene er ikke oplyst, men den antages at være lille (variationskoefficient skønnet til 1%), dog noget større for palmeolie og sukkerroer (variationskoefficient skønnet til 10%), da udbytterne for de mellemliggende forarbejdningsled ikke er verificeret.

Tabel 5.2.2
Afgrøder til 1 kg rugbrød.

Afgrøde	kg pr. kg rugbrød	Variationskoefficient
Rug	0,536	1%
Sukkerroe	0,286	10%
Hvede	0,122	1%
Hørfrø, U.K.	0,015	1%
Hørfrø, Rusland	0,015	1%
Palmefrugt	0,032	10%
Sesamfrø	0,004	1%
Solsikkefrø	0,030	1%
Ialt	1,040	1%

Høstudbytte pr. ha af de forskellige afgrøder

Høstudbyttet pr. ha er beregnet som gennemsnit for flere år ud fra de nationale statistikker fra oprindelseslandet undtagen for palmefrugt, hvor der er brugt specifikke data fra en Malaysisk plantage i 1990.

Usikkerheden på høstudbyttet er beregnet ved at tage tal for flere år og betragte de enkelte data som uafhængige. Beregningerne er gengivet i tabel 5.2.3.

Tabel 5.2.3
Beregnet usikkerhed for udbytte pr. ha af de enkelte afgrøder.

År Afgrøde Land	Gennemsnitligt udbytte i kg pr. ha						
	Rug Danmark	Sukkerroe Danmark	Hvede Danmark	Hørfrø U.K.	Hørfrø Rusland	Sesamfrø Venezuela	Solsikkefrø U.S.A.
1979-81	-	-	-	-	-	183	667
1990	3500	45600	6230	1751	232	500	1378
1991	4910	49900	7090	1727	206	529	1515
1992	4970	53300	7430	1571	214	529	1163
Antal data	3	3	3	3	4	4	4
Sum	13380	148800	20750	5049	835	2225	5379
Middelværdi	4460	49600	6917	1683	209	556	1345
Varians	692100	14890000	382533	9552	413	5638	21199
Standardafvigelse	832	3859	618	98	20	75	146
Variationskoefficient	18,7%	7,8%	8,9%	5,8%	9,7%	13,5%	10,8%

Datakvaliteten er god bortset fra nationale gennemsnit, som ikke nødvendigvis svarer til forholdene på de konkrete landbrugsbedrifter, der leverer ingredienserne til rugbrødet. Den yderligere usikkerhed, der kan stamme herfra, afhænger af variationen mellem forskellige landbrugsbedrifter og hvor stor en andel af den nationale afgrøde, der afsættes til den undersøgte rugbrødsproduktion. Selv for den væsentligste ingrediens (rug) er denne andel lille (ca. 7%), og den er derfor anvendt som et estimat af usikkerheden på variationen i udbytte mel-

lem de forskellige danske amter. I 1992 var amts-variationen udtrykt som variationskoefficenter: 17,2% for rug, 9,6% for sukkerroe og 9,9% for hvede. Da tilsvarende variationer ikke kendes for de udenlandske afgrøder, anvendes her det danske tal for rug (CV: 17,2%) som et skøn for variationen mellem de enkelte landbrugsbedrifter.

Kombineret med usikkerhederne fra tabel 5.2.3 fås samlede variationskoefficenter som vist i tabel 5.2.4.

Tabel 5.2.4

Samlet usikkerhed for høstudbytte pr. ha.

Afgrøde Land	Rug DK	Sukkerroe DK	Hvede DK	Hørfrø U.K.	Hørfrø Rusland	Sesamfrø Venezuela	Solsikkefrø U.S.A.
Års-variation (CV)	18,7%	7,8%	8,9%	5,8%	9,7%	13,5%	10,8%
Geografisk variation (CV)	17,2%	9,6%	9,9%	17,2%	17,2%	17,2%	17,2%
Samlet variations- koefficient	25%	12%	13%	18%	20%	22%	20%

For palmefrugt er udbyttet 22.500 kg pr. ha og variationskoefficienten skønnes til 25%, idet det ikke vides om Malaysia er det faktiske eksportland.

Forbrug af pesticider pr. ha

Pesticidforbruget for de enkelte afgrøder er ikke kendt. I stedet er anvendt det totale forbrug af pesticider i Danmark divideret med det totale dyrkede areal. Der er brugt tal fra 1988-92. Beregningerne fremgår af tabel 5.2.5.

Tabel 5.2.5

Forbrug af pesticider pr. ha opgjort i kg virksomt stof.

År	Total forbrug kg	Dyrket areal ha	Forbrug pr. ha kg/ha
1988	5.521.000	2.786.603	1.981
1989	6.098.000	2.774.127	2.198
1990	5.461.000	2.788.276	1.959
1991	5.348.000	2.769.657	1.931
1992	5.358.000	2.756.327	1.944
Antal data			5
Sum			10.013
Middelværdi			2.003
Varians			0.0123065
Standardafvigelse			0.1109348
Variationskoefficient			5,5

Kvaliteten af disse data kan angives som 2 kg/ha (CV: 5,5 %) (2,1,1,5,4). Begrundelsen for datakvalitetsindekset gives i tabel 5.2.5.

Tabel 5.2.6
Datakvalitetsindeks for pesticidforbrug for danske afgrøder.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Variation mellem år		CV: 5,5%
Pålidelighed	Data tager ikke højde for lagersforskydninger og der indgår forbrug i private haver.	2	CV: 2,5%, estimeret
Fuldstændighed	Perioden 1988-92 er tilstrækkelig til at dække variationer mellem år.	1	-
Tidsmæssig korrelation	Data er dækkende for 1992.	1	-
Geografisk korrelation	Gennemsnitlige tal for Danmark, ikke for de enkelte leverende gårde.	2	CV: 16,5%, estimeret
Teknologisk korrelation	Gennemsnit for samtlige afgrøder, ikke for de faktiske afgrøder.	4	CV: 33%, estimeret
Samlet usikkerhed			CV: 37%

Tages der højde for de supplerende usikkerheder, kan forbruget af pesticider altså alternativt skrives 2 kg/ha (CV: 37%)(m,1,1,m,m).

Da der mangler tal for pesticidsforbruget til udenlandske afgrøder, er de danske tal (2 kg/ha) også anvendt for de importerede afgrøder. Som vist i tabel 5.2.7 giver dette en dårligere score på de sidste 2 indikatorer (geografisk og teknologisk korrelation) og dermed en forøget usikkerhed.

Tabel 5.2.7
Datakvalitetsindeks for pesticidforbrug for udenlandske afgrøder.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Variationer mellem år		CV: 5,5%
Pålidelighed	Data tager ikke højde for lagersforskydninger og der indgår forbrug i private haver.	2	CV: 2,5%, estimeret
Fuldstændighed	Perioden 1988-92 er tilstrækkelig til at dække variationer mellem år.	1	-
Tidsmæssig korrelation	Data er dækkende for 1992.	1	-
Geografisk korrelation	Gennemsnitlige tal for Danmark, ikke for op-rindelseslandet.	5	CV: 50%, estimeret
Teknologisk korrelation	Gennemsnit for samtlige afgrøder, ikke for de faktiske afgrøder. Sprøjte-teknologien i andre lande kan også være væsentligt forskellig fra danske forhold.	5	CV: 50%, estimeret
Samlet usikkerhed			CV: 71%

Energiforbrug til produktion af pesticider

Der anvendes et gennemsnit for energiforbrug til 39 specifikke pesticider opgivet i en ældre kilde. Energiforbruget kan angives som 240 MJ/kg (CV: 58,7%) (3,1,5,3,4). Begrundelsen for datakvalitetsindekset gives i tabel 5.2.8. Tages højde for denne datakvalitet - jf. tabel 5.2.8 - forøges middelværdien med 20% og varianskoefficienten øges til 62%. Middelværdi og datakvalitetsindeks er hermed 288 MJ/kg (CV: 62%)(m,1,m,m,m).

Tabel 5.2.8**Datakvalitetsindeks for energiforbrug til pesticid fremstilling.**

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Beregnet variation mellem de 39 pesticider i kilden.		CV: 58,7%
Pålidelighed	Kilden anvender en computermodel baseret delvist på procesbeskrivelser fra patentansøgninger, ikke på målte data.	3	CV: 16,5%, estimeret
Fuldstændighed	Data er repræsentative for de 39 nævnte pesticider.	1	-
Tidsmæssig korrelation	Data er fra 1976 delvist baseret på ældre data. Det vurderes at energieffektiviteten er steget siden.	5	Reducér middelværdi med 40%, estimeret
Geografisk korrelation	Generelle tal er anvendt for specifikke pesticider anvendt i Danmark. Forskellen i energiforbrug mellem forskellige geografiske produktionsteder vurderes dog at være begrænset.	3	CV: 10%, estimeret
Teknologisk korrelation	Middelværdien ændres for at tage højde for det større energiforbrug til moderne, komplexe pesticider (f.eks. Pyretroider) sammenlignet med ældre, mere simple forbindelser.	4	Forøg middelværdi med 60%, estimeret.
Samlet usikkerhed			CV: 62% og middelværdi forøges med 20%.

Forbrug af N, P og K kunstgødning pr. ha pr. dansk afgrøde
Kunstgødningsforbruget pr. afgrøde (F_a) er beregnet som:

$$F_a = B_a D \quad (13)$$

hvor B_a er den anbefalede gødningsmængde pr. ha for den pågældende afgrøde. D er den gennemsnitlige dækningsgrad (D) for kunstgødning beregnet som:

$$D = \frac{F_{tot}}{B_{tot}} \quad (14)$$

hvor F_{tot} er det samlede danske kunstgødningsforbrug. B_{tot} er det samlede danske gødningsbehov beregnet som:

$$B_{tot} = \sum_{a=1}^n B_a A_a \quad (15)$$

hvor A_a er det samlede danske areal for afgrøden a.

Herved allokeres samlede mængde kunstgødning til en afgrøde proportionalt med gødningsbehovet for den pågældende afgrøde. Der tages således ikke højde for, at visse afgrøder tildeles forholdsmaessigt mere husdyrgødning end andre. Usikkerheden på dette allokeringsprincip kan medregnes som en supplerende usikkerhed på D_N , idet denne altså i praksis varierer mellem afgrøder. Dette er imidlertid ikke gjort i nærværende studie.

Benyttelsen af den anbefalede gødningsmængde til en afgrøde (B_a), som et udtryk for den faktiske gødningsmængde, kan gives datakvali-

tetsindekset (2,1,1,2,1), hvilket kan fortolkes til en yderligere usikkerhed på 7% (m,1,1,m,1). Dette begrundes i tabel 5.2.9.

Tabel 5.2.9

Datakvalitetsindeks for den anvendte kunstgødningsmængde for danske afgrøder (B_a).

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Anbefalingerne angivet uden usikkerhed.		0%
Pålidelighed	Fordelingen på afgrøder kan godt i praksis være forskellig fra det anbefalede.	2	CV: 5%, estimeret
Fuldstændighed	Anbefalingerne dækker samtlige danske afgrøder.	1	-
Tidsmæssig korrelation	Anbefalingerne dækker det aktuelle år.	1	-
Geografisk korrelation	Anbefalingerne er gennemsnit og svarer ikke nødvendigvis til de lokale forhold på de levende gårde.	2	CV: 5%, estimeret
Teknologisk korrelation	Anbefalingerne dækker de aktuelle afgrøder og den aktuelle teknologi.	1	-
Samlet usikkerhed			CV: 7%

Datakvalitetsindekset for det totale kunstgødningsforbrug F_{tot} er (2,1,1,1,1), idet tallet ikke er justeret for lagerforskydninger og inkluderer forbruget i private haver. Den afledte usikkerhed herved kan skønnes til en variationskoefficient på 2,5% (m,1,1,1,1). Arealet med forskellige afgrøder (A_a) er af Danmarks Statistik beregnet med en variationskoefficient på 0,3% (1,1,1,1,1). Denne usikkerhed kan overføres direkte til det samlede danske gødningsbehov (B_{tot}).

Foretages beregningen af dækningsgraden (D) for flere år, kan variationen over tid beregnes. I det oprindelige baggrundsmateriale til Weidema *et al.* (1995) er dækningsgraden for kvælstof-kunstgødning D_N således beregnet for 1990 og 1992 til henholdsvis 97% og 90% (se tabel 5.2.10 for beregningen for 1992). Forskellen mellem 1990 og 1992 er imidlertid ikke nødvendigvis udtryk alene for variationen mellem år, da der på grund af tiltagende offentlig regulering af håndteringen af husdyrgødning sker en stadig bedre udnyttelse af husdyrgødningen fra år til år. Derved bliver det totale kunstgødningsforbrug (F_{tot}) og dermed dækningsgraden (D) lavere fra år til år. Denne forskydning bør principielt elimineres ved at foretage en regression af beregningen for flere år, hvorved den rene års-variation kan isoleres. En sådan regression er imidlertid ikke gennemført i nærværende studie, og tallene for de to år (1990 og 1992) bruges dersør som ligeværdige udtræk til beregning af variationskoefficienten på års-variationen af D_N ; denne kan beregnes til 5,3%. Denne værdi bruges generelt for D.

Tabel 5.2.10

Beregning af gennemsnitlig dækningsgrad med kunstgødning i dr 1992.

Danske afgrøder	Anbefalet gødningsmængde pr. ha 1992			Areal 1992	Anbefalet gødningsmængde 1992		
	N	P	K		N	P	K
	kg pr. ha	1000 ha	1000 kg				
Vinterhvede	180	25	60	567	102060	14175	34020
Vårhvede	180	25	60	15	2700	375	900
Rug	140	20	50	88	12320	1760	4400
Vinterbyg	170	25	60	151	25670	3775	9060
Vårbyg	130	25	60	759	98670	18975	45540
Havre	130	25	60	28	3640	700	1680
Blandsæd	130	25	60	3	390	75	180
Bælgplanter	0	20	70	118	0	2360	8260
Læggekartofler	140	30	125	9	1260	270	1125
Industri-kartofler	140	30	125	31	4340	930	3875
Spisekartofler	160	30	150	13	2080	390	1950
Sukkerroer	120	35	150	65	7800	2275	9750
Sukkerroer til foder	180	35	200	79	14220	2765	15800
Foderrør	160	45	200	1	160	45	200
Vinternraps	170	25	100	118	20060	2950	11800
Vårraps	130	20	90	63	8190	1260	5670
Hør	120	20	80	1	120	20	80
Andre industrifrø	120	20	80	0	0	0	0
Frø til udsæd	120	20	60	52	6240	1040	3120
Lucerne	0	40	250	11	0	440	2750
Majs til opfodring	160	35	175	20	3200	700	3500
Korn til helsæd	130	40	170	64	8320	2560	10880
Andet grønfoder	130	40	170	3	390	120	510
Græs og kløvergræs	300	40	200	255	76500	10200	51000
Havebrugsafgrøder	400	40	200	29	11600	1160	5800
Øvrige arealer i omdriften	0	0	0	4	0	0	0
Permanent græs	0	0	0	208	0	0	0
I alt	-	-	-	2756	409930	69320	231850
Kunstgødningsforbrug 1992					369500	33200	111800
Andel kunstgødning (D) af samlet gødningsbehov					90%	48%	48%

Forbrug af N, P og K kunstgødning pr. ha pr. udenlandsk afgrøde
 For de fleste udenlandske afgrøder kendes de anbefalede gødningsmængder ikke, og en allokering af kunstgødningsmængden på dette grundlag kan derfor ikke finde sted. I stedet bruges blot det gennemsnitlige forbrug af kunstgødning pr. ha i oprindelseslandet. I tabel 5.2.11 ses dette beregnet for kvælstof.

Tabel 5.2.11

Gennemsnitligt forbrug af N-kunstgødning til visse udenlandske afgrøder.

	Forbrug af kvælstof-kunstgødning i oprindelseslandet i alt (ton pr. år):			
	U.K.	Rusland	Venezuela	U.S.A.
90/91 (regnes til 1991)	1525000	8668700	205000	10239255
85/86 (regnes til 1986)	1568000	10950000	215485	9456852
80/81 (regnes til 1981)	1240000	8262000	113029	10817039
	Opdyrket areal i oprindelseslandet i alt (1000 ha):			
	U.K.	Rusland	Venezuela	U.S.A.
1991	6552	224400	3210	185742
1986	6958	227399	3130	187765
1981	6920	226600	3080	188755
	Forbrug af kunstgødning (kg pr. ha pr. år):			
	U.K.	Rusland	Venezuela	U.S.A.
1991	232.8	38.6	63.9	55.1
1986	225.4	48.2	68.8	50.4
1981	179.2	36.5	36.7	57.3
Statistiske data:				
Antal data	3	3	3	3
Sum	637.3	123.2	169.4	162.8
Middelværdi	212.4	41.1	56.5	54.3
Varians	842.4	38.7	299.4	12.6
Standardafvigelse	29.025	6.220	17.302	3.550
Variationskoefficient	13.7%	15.1%	30.6%	6.5%

Variationskoefficienterne i tabel 5.2.11 er til dels påvirket af, at store lande (f.eks. USA) vil have en mindre relativ variation mellem de enkelte år end små lande (f.eks. Venezuela). Variationskoefficienten afspejler således ikke de egentlige forskelle mellem afgrøderne, hvilket også fremgår af datakvalitetsindeksset (2,1,1,2,4), jf. tabel 5.2.12.

Tabel 5.2.12

Datakvalitetsindeks for forbrug af kunstgødning pr. ha til visse udenlandske afgrøder.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Pålidelighed	Data tager ikke højde for lagerforskydninger og privat forbrug.	2	CV: 2,5%, estimeret
Fuldständighed	Data dækker al kunstgødning.	1	-
Tidsmæssig korrelation	Anbefalingerne dækker det aktuelle år.	1	-
Geografisk korrelation	Gennemsnitsal svarer ikke nødvendigvis til de lokale forhold på de leverende gårde.	2	CV: 5%, estimeret
Teknologisk korrelation	Data er gennemsnitsal og gødningsmængderne for de aktuelle afgrøder kan variere meget.	4	CV: 50%, estimeret

Sammenlagt bliver variationskoefficienterne da 52,1% for U.K. (hørfrø), 52,5% for Rusland (hørfrø), 58,9% for Venezuela (sesamfrø) og 50,7% for U.S.A. (solsikkekerner).

Som det ses af ovenstående, domineres den samlede usikkerhed helt af usikkerheden på den teknologiske korrelation. Da dette også vil være tilfældet for P og K gødning, er der ikke foretaget en særskilt beregning af variationskoefficienterne for disse gødninger. I de videre beregninger i afsnit 6.2. benyttes samme variationskoefficienter som for kvælstofgødning.

Ud over de nævnte afgrøder importeres palmeolie fra Malaysia. Data for kunstgødningsforbruget kendes her specifikt for palmefrugt. For kvælstofgødning angiver kilden gødningstypen urea og et interval på 70 - 140 kg/ha. Dette fortolkes som et 95% konfidensinterval, der antages at være repræsentativt for dyrkning af palmefrugt generelt.

Variationskoefficienten bliver da 17%. Kilden angiver intervaller af samme størrelse for P og K gødning.

Energiforbrug til produktion af N kunstgødning

Den originale kilde angiver energiforbruget til fremstilling af kvælstofgødning som et interval for de enkelte processer i fremstillingen.

For ammoniumnitrat og komplekse gødninger opgives intervallerne i tabel 5.2.11 med et samlet interval fra 25 MJ/kg til 67 MJ/kg, der antages at være repræsentativt for fabrikker af meget forskellig alder og dermed teknologi. Betragtes de oplyste intervaller som 95% konfidensintervaller for normalfordelte data fås en middelværdi på 46 MJ/kg og en variationskoefficient på 16%, jf. tabel 5.2.13.

Tabel 5.2.13

Energiforbrug til fremstilling af ammoniumnitrat og komplekse gødninger.

MJ/kg N til	Oplyst interval	Middelværdi og standardafvigelse i et antaget 95% konfidensinterval
Ammoniak-fremstilling	35 til 60	47,5 (standardafvigelse 6,75)
Ammoniak til salpetersyre	-12 til -8	-10 (standardafvigelse 1)
Yderligere processer	0 til 5	2,5 (standardafvigelse 1,25)
Granulering	2 til 10	6 (standardafvigelse 2)
I alt	25 til 67	46 (CV: 16%)

De tilsvarende intervaller for urea er vist i tabel 5.2.14, og givet med de samme antagelser 63 MJ/kg (CV: 11%). Denne værdi anvendes for N-gødning til palmefrugt.

Tabel 5.2.14

Energiforbrug til fremstilling af urea.

MJ/kg N til	Oplyst interval	Middelværdi og standardafvigelse i et antaget 95% konfidensinterval
Ammoniak-fremstilling	35 til 60	47,5 (standardafvigelse 6,75)
Ammoniak til fast urea	8 til 17	12,5 (standardafvigelse 2,25)
Granulering	1 til 5	3 (standardafvigelse 1)
I alt	44 til 82	63 (CV: 11%)

I det undersøgte studie (Weidema *et al.*, 1995) er der for dansk N-kunstgødning anvendt "bedste skøn" for de enkelte processer i stedet for de oprindelige intervaller, idet det er fundet godtgjort, at leverandørerne til det danske marked i gennemsnit anvender en teknologi i den bedre ende af intervallet. Denne antagelse kan fortolkes som en indsnævring af de ovennævnte intervaller, idet der introduceres et (begrænset) kendskab til den anvendte teknologi. De nye standardafvigelser vises i tabel 5.2.15, hvoraf fås en værdi på 38 MJ/kg (CV: 10%) for N-kunstgødning anvendt i Danmark.

Tabel 5.2.15*Energiforbrug til fremstilling af N-kunstgødning anvendt i Danmark.*

MJ/kg N	Estimeret interval	Middelværdi og standardafvigelse i et antaget 95% konfidensinterval
Ammoniak-fremstilling	35 til 47	41 (standardafvigelse 3)
Ammoniak til salpetersyre	-12 til -8	-10 (standardafvigelse 1)
Yderligere processer	0 til 2	1 (standardafvigelse 0,5)
Granulering	2 til 10	6 (standardafvigelse 2)
	25 til 51	38 (CV: 10%)

Fordelingen mellem urea og andre N-gødninger er meget afhængig af oprindelseslandet, idet udviklingslande typisk bruger en meget større andel af urea end vestlige lande (se tabel 5.2.16). Variationskoefficienten på disse fordelinger er størst i de mindste lande.

Da fordelingen mellem urea og de øvrige gødningstyper ikke findes oplyst for alle lande, antager Weidema *et al.* (1995) som en generel worst-case antagelse en ligelig fordeling mellem urea og ammonium-nitrat/komplekse gødninger. Anvendes de ovenfor nævnte middelværdier på 46 MJ/kg og 63 MJ/kg, fås altså en middelværdi for N-gødning anvendt til importerede afgrøder på 54,5 MJ/kg. En mere korrekt procedure ville have været at anvende de kendte fordelinger for de enkelte lande, idet det kun er for Rusland fordelingen ikke kendes (se tabel 5.2.16). Et rimeligt estimat for Rusland baseret på fordelingen i de øvrige lande er 25% urea (CV: 20%). Denne korrigerede procedure er anvendt i analysen i kapitel 6.2.

Tabel 5.2.16*Fordeling mellem urea og andre kvælstofgødninger beregnet for 1987-1990.*

	Urea's andel af total N	Usikkerhed på fordelingen
U.K. (Hørfrø)	9%	CV: 10%
Rusland (Hørfrø)	25%, estimeret	CV: 20%, estimeret
Venezuela (Sesamfrø)	50%	CV: 20%
U.S.A. (Solsikkefrø)	15%	CV: 4%

Energiforbrug til produktion af P kunstgødning

Den originale kilde angiver energiforbruget til fremstilling af phosphorgødning opdelt på enkeltprocesser med intervaller som vist i tabel 5.2.17, der antages at være repræsentativt for fabrikker af meget forskellig alder og dermed teknologi. Betragtes de oplyste intervaller som 95% konfidensintervaller for normalfordelte data, fås en middelværdi på 28,3 MJ/kg og en variationskoefficient på 15%, jf. tabel 5.2.17.

Tabel 5.2.17*Energiforbrug til fremstilling af phosphorgødning.*

MJ/kg P i gødning	Oplyst interval	Middelværdi og standardafvigelse i et antaget 95% konfidensinterval
Brydning	6,4 til 12,8	9,6 (standardafvigelse 1,6)
Produktion af svovlsyre	-8,1 til 6,8	-1,3 (standardafvigelse 3,4)
Produktion af phosphorsyre	9,5 til 16,7	13,1 (standardafvigelse 1,8)
Produktion af superfosfater	5,7 til 8,0	6,9 (standardafvigelse 0,6)
	13,5 til 44,3	28,3 (CV: 15%)

Energiforbrug til produktion af K kunstgødning
Der anvendes et tal opgivet fra Norsk Hydro på 5 MJ/kg K.

Kvaliteten af data kan angives som 5 MJ/kg (4,2,1,3,3). Begrundelsen for datakvalitetsindekset gives i tabel 5.2.18.

Tabel 5.2.18
Datakvalitetsindeks for energiforbrug til produktion af K kunstgødning.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindeelig usikkerhed	Ikke oplyst.		-
Pålidelighed	Data må betragtes som et kvalificeret estimat, da der ikke er oplyst andet.	4	CV: 10%, estimeret
Fuldstændighed	Data oplyst af Norsk Hydro, dvs. fra én enkelt producent.	2	CV: 10%, estimeret
Tidsmæssig korrelation	Data fra 1990.	1	-
Geografisk korrelation	Data formodentlig for norske forhold. Geografisk variation på energiforbrug er dog antaget at være lille.	3	CV: 5%, estimeret
Teknologisk korrelation	Kilden angiver at med en ældre proces (krystallisation) kan energiforbruget være op til 50% større.	3	CV: 25%
Samlet usikkerhed			CV: 29%

Forbrug af kalk pr. ha pr. afgrøde

De anvendte mængder kalk er beregnet som det årlige gennemsnitsforbrug i Danmark i årene 1983 til 1992 (1,4 mio. kg CaCO₃) allokeret til 1992 arealet med de afgrøder, der er særligt kalktrængende (beder, byg, hvede og bælgplanter), i alt 1.766.000 ha.

Data kan angives som 790 kg/ha (2,2,1,3,3). Datakvalitetsindekset kan fortolkes som en usikkerhed på 10%, jf. tabel 5.2.19.

Tabel 5.2.19
Datakvalitetsindeks for forbrug af kalk.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Ikke oplyst		-
Pålidelighed	Data er baseret på oplysninger fra leverandører.	2	CV: 2,5%, estimeret
Fuldstændighed	Arealet med forskellige afgrøder er bestemt på basis af stikprøver. Usikkerheden herved er opgivet i den primære kilde.	2	CV: 0,3%, estimeret
Tidsmæssig korrelation	Data dækker det aktuelle år.	1	-
Geografisk korrelation	Gennemsnit for Danmark, ikke for de leverende gårde.	2	CV: 10%, estimeret
Teknologisk korrelation	Data dækker den aktuelle teknologi.	1	-
Samlet usikkerhed			CV: 10,3%

Energiforbrug til fremstilling af kalk

Energiforbruget er et estimat på baggrund af et review af flere kilder. Data kan angives som 0,8 MJ/kg (5,5,5,5,3). Begrundelsen for datakvalitetsindexet gives i tabel 5.2.20.

Tabel 5.2.20*Datakvalitetsindeks for energiforbrug til fremstilling af kalk.*

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Ikke oplyst (se pålidelighed)	-	-
Pålidelighed	Estimat på basis af litteratur-review, hvor der er fundet værdier op til 2,35 MJ/kg. Den opgivne usikkerhed fås ved at antage at denne værdi er den øvre grænse for et 95% konfidenstinterval for middelværdien 0,8 MJ/kg.	5	CV: 97%
Fuldstændighed	Repræsentativiteten af den primære kilde er ukendt.	5	CV: 25%, estimeret
Tidsmæssig korrelation	Kilden benytter ældre data, men det antages at der delvist er taget højde herfor i estimatet.	5	CV: 6%, estimeret
Geografisk korrelation	Den geografiske dækning er ukendt.	5	CV: 10%, estimeret
Teknologisk korrelation	Data for kalk, men teknologi ukendt.	3	CV: 25%, estimeret
Samlet usikkerhed			CV: 104%

Tages der højde for den afledte usikkerhed, kan energiforbruget til fremstilling af kalk angives som 0,8 MJ/kg (CV: 104%) (m,m,m,m,m).

Forbrug af husdyrgødning pr. ha pr. afgrøde

Det antages, at der kun anvendes husdyrgødning til de danske afgrøder. Den samlede mængde husdyrgødning i Danmark ($H_{tot} = 45.800$ mio. kg) skal allokeres mellem husdyrproduktionen (i det omfang den ikke har nogen økonomisk værdi) og planteproduktionen. Den sidstnævnte mængde (H_{pl}) beregnes som den mængde, der er nødvendig til at dække forskellen mellem forbruget af N i kunstgødning ($F_{tot,N} = 369,5$ mio. kg) og de anbefalede mængder N for samtlige danske afgrøder ($B_{tot,N} = 410$ mio. kg) ved en udnyttelsesgrad på 40% af N i husdyrgødning, altså:

$$H_{pl} = ((B_{tot,N} - F_{tot,N})/0,4)/(H_{tot,N}/H_{tot}) \quad (16)$$

hvor $H_{tot,N}$ (= 246 mio. kg) er mængden af kvælstof i den totale mængde husdyrgødning. Anvendes samme terminologi som i afsnittet om forbrug af kunstgødning, kan $B_{tot,N} - F_{tot,N}$ også udtrykkes som $1 - D_N$.

Usikkerheden på B_{tot} , F_{tot} og D_N er allerede beskrevet under forbrug af kunstgødning. $H_{tot,N}$ og H_{tot} er kvalificerede estimeret med estimeret variationskoefficient på 10%.

For at få H_{pl} udtrykt pr. ha, skal mængden af husdyrgødning allokeret til planteproduktion divideres med det totale dyrkede areal i Danmark (2.756.000 ha). Det antages herved, at husdyrgødningen fordeles ligeligt på alle afgrøder i forhold til N behovet. Denne antagelse modsvarer antagelsen om, at kunstgødningsforbruget allokeres proportionalt med gødningsbehovet for de enkelte afgrøder. Som tidligere nævnt kan usikkerheden på denne antagelse medregnes som en supplerende usikkerhed på D_N , idet denne altså i praksis varierer mellem afgrøder (men dette er ikke gjort i dette studie).

Energiforbrug til fremstilling af husdyrgødning

Energiforbruget er beregnet som elforbruget pr. slagtesvin divideret med gødningsmængden pr. slagtesvin. Elforbruget er ud fra kapacitetsomkostninger, som haves fra to forskellige kilder (Pedersen *et al.*, 1993; Porskrog *et al.*, 1993), beregnet til 35 kWh/slagtesvin (CV: 25%). Dette svarer meget godt til det interval, der opgives i "Håndbog for svinehold 1993" (Bomholt *et al.*, 1992). Gødningsmængden bygger delvis på kvalificerede skøn, hvortil vi estimerer en samlet variationskoefficient således, at gødningsmængden pr. slagtesvin kan angives som 1002 kg (CV: 10%).

Det energiforbrug, der beregnes på denne måde, skal yderligere tilføjes usikkerhed som følge af, at disse danske gennemsnitstal for svin anvendes for alle typer husdyrgødning (score 4 for teknologisk korrelation estimeret til en CV på 25%) og for de konkrete gårde (score 2 for geografisk korrelation estimeret til en CV på 5%).

Energiforbrug pr. ha til markarbejde for danske afgrøder

Energiforbruget pr. ha er beregnet ud fra brændstofforbruget for udvalgte maskiner målt over flere år ganget med det målte timesforbrug til forskellige markarbejder i forskellige standardscenarier. For de maskintyper, hvor brændstofforbruget ikke er målt i praksis, er det antaget at svare til forbruget ved 80% af maksimal belastning. Ved anvendelse af tallene er energiforbruget til høstprocesserne justeret proportionalt med det aktuelt anvendte høstudbytte. Spredning af husdyrgødning er antaget at kræve samme energiforbrug som spredning af kunstgødning.

De anvendte tal er 2309 MJ/ha for rug, 4186 MJ/ha for sukkerroer og 2836 MJ/ha for hvede. Usikkerheden på de målte tal er meget varierende mellem redskabs- og maskintyperne med variationskoefficienter fra 10 til 63%, størst for de mindste operationer. For de samlede energiforbrug antages variationskoefficienten at være i den lave ende af dette interval (estimeret til 20%), da de domineres af enkelte store poster.

Datakvaliteten kan angives som (2,1,2,2,1), hvoraf afledes en reduktion i middelværdien på 10% og samlet variationskoefficient på 34%, hvilket begrundes i tabel 5.2.21.

Tabel 5.2.21
Datakvalitetsindeks for energiforbrug til markarbejde for danske afgrøder.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Måleusikkerhed		CV: 20%
Pålidelighed	Generelt gode data med enkelte antagelser.	2	CV: 2,5%, estimeret
Fuldstændighed	Data dækker alt markarbejde.	1	-
Tidsmæssig korrelation	Data fra slutningen af 1980'erne.	2	Reducér middel- værdi med 10%, esti- meret
Geografisk korrelation	Danske data, ikke fra de konkrete leverende gårde.	2	CV: 10%, estimeret
Teknologisk korrelation	Data for aktuelle afgrøder og aktuel teknologi.	1	
Samlet usikkerhed			CV: 23% og middel- værdi reduceres med 10%

De korrigerede værdier er 2078 MJ/ha for rug, 3767 MJ/ha for sukkerroe og 2552 MJ/ha for hvede, alle med datakvaliteten CV: 23% (m,1,m,m,1).

Energiforbrug til markarbejde pr. ha for udenlandske afgrøder.

For de udenlandske afgrøder findes kun et generelt tal for industrielle afgrøder fra Grækenland. Ved anvendelse af tallet er energiforbruget til høstprocesserne justeret proportionalt med høstudbyttet for den aktuelle afgrøde.

Datakvaliteten kan angives som (3,1,2,5,5), hvoraf afledes en reduktion i middelværdien på 10% og samlet variationskoefficient på 52%, hvilket begrundes i tabel 5.2.22. De resulterende værdier er opført i tabel 5.2.23.

Tabel 5.2.22
Datakvalitetsindeks for energiforbrug til markarbejde for udenlandske afgrøder.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Måleusikkerhed som for danske afgrøder.		CV: 20%
Pålidelighed	Der mangler oplysninger om antagelser og verifikation.	3	CV: 10%, estimeret
Fuldstændighed	Data dækker alt markarbejde.	1	-
Tidsmæssig korrelation	Data fra slutningen af 1980'erne. Det vurderes at energi-effektiviteten er steget siden.	2	Reducér middel- værdi med 10%, esti- meret
Geografisk korrelation	Græske data. Gennemsnitlig variation på data for samme afgrøde fra forskellige lande bereg- net for kartofler, sukkerroe og majs.	5	CV: 40%
Teknologisk korrelation	Data for anden afgrøde med delvis anden tek- nologi. Variation beregnet ud fra 8 forskellige scenarier for dansk byg.	5	CV: 25%
Samlet usikkerhed			CV: 52% og middel- værdi reduceres med 10%

Tabel 5.2.23*Energiforbrug til markarbejde for udenlandske afgrøder.*

MJ pr. ha	Oprindelig værdi	Korrigeret værdi
Hørfrø, U.K.	1761	1585
Hørfrø, Rusland	1450	1305
Sesamfrø	1488	1339
Solsikkefrø	1675	1508
Datakvalitet	(CV: 20%)(3,1,2,5,5)	(CV: 52%)(m,1,m,m,m)

For palmefrugt, hvor der er tale om målte tal fra 1990, kan energiforbruget angives som 2154 MJ/ha (CV: 20%)(1,1,1,1,1), idet der regnes med samme oprindelige usikkerhed som for de danske tal.

Energiforbrug til tørring af korn-afgrøder

Det antages, at kun rug og hvede skal tørres fra 18% til 14% vandindhold.

Energiforbruget til tørring kan angives som 0,24 MJ pr. kg korn (4,1,2,1,1). Af datakvaliteten afledes en reduktion i middelværdien på 10% og samlet variationskoefficient på 25%, hvilket begrundes i tabel 5.2.24.

Tabel 5.2.24*Datakvalitetsindeks for energiforbrug til tørring af korn.*

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed	Ikke oplyst.	-	-
Pålidelighed	Kvalificeret estimat. Usikkerhed på antagelsen om, at alt korn i gennemsnit skal tørres 4% ned.	4	CV: 25%, estimeret
Fuldständighed	Data dækker alt korn.	1	-
Tidsmæssig korrelation	Data fra slutningen af 1980'erne. Det vurderes, at energi-effektiviteten er steget siden.	2	Reducér middelværdi med 10%, estimeret
Geografisk korrelation	Danske data.	1	-
Teknologisk korrelation	Data for aktuel teknologi.	1	-
Samlet usikkerhed			CV: 25% og middelværdi reduceres med 10%

Den resulterende værdi er altså 0,22 MJ/kg (CV: 25%)(m,1,m,1,1).

5.2.3 Samlet vurdering af datakvaliteten

Datakvaliteten for de indgående parametre er sammenfattet i tabel 5.2.25 (sammenlign med tabel 5.2.1).

Tabel 5.2.25

Middelværdi, variationskoefficient, datakvalitetsindeks og korrigerede variationskoefficient for de indgående data.

Data	Middelværdi (μ)	Oprindelig CV	Datakvali- tetsindeks	Korrigerede CV: (og evt. μ)
Afgrøder til 1 kg rugbrød - sukkerroe og palmefrugt - øvrige afgrøder	Se tabel 5.2.2 Se tabel 5.2.2	1% 1%	(2.1.1.1.1) (1.1.1.1.1)	10% 1%
Høstudsbytte pr. ha af: - rug - sukkerroe - hvede - hørfrø, U.K. - hørfrø, Rusland - palmefrugt - sesamfrø - solsikkekerner	4460 kg/ha 49600 kg/ha 6917 kg/ha 1683 kg/ha 209 kg/ha 22500 kg/ha 556 kg/ha 1345 kg/ha	19% 8% 9% 6% 10% - 14% 11%	(1.1.1.2.1) (1.1.1.2.1) (1.1.1.2.1) (1.1.1.2.1) (1.1.1.2.1) (1.1.1.3.1) (1.1.1.2.1) (1.1.1.2.1)	25% 12% 13% 18% 20% 25% 22% 20%
Forbrug af pesticider pr. ha - for danske afgrøder - for udenlandske afgrøder	2 kg/ha 2 kg/ha	6% 6%	(2.1.1.2.4) (2.1.1.5.5)	37% 71%
Energiforbrug til produktion af pes-ticider	240 MJ/kg	59%	(3.1.5.3.4)	62% (μ = 288 MJ/kg)
Forbrug af N, P og K kunstgødning pr. ha for danske afgrøder beregnes ud fra: - anbefalede gødningsmængder - totalt kunstgødningsforbrug - areal af den enkelte afgrøde	Se tabel 5.2.10 Se tabel 5.2.10 Se tabel 5.2.10	0% - 0.3%	(2.1.1.2.1) (2.1.1.1.1) (1.1.1.1.1)	7% 2.5% 0.3%
Forbrug af N kunstgødning pr. ha for udenlandske afgrøder: - hørfrø, U.K. - hørfrø, Rusland - palmeolie - sesamfrø - solsikkekerner (for P og K anvendes samme varia-tionskoefficiente)	212 kg/ha 41 kg/ha 105 kg/ha 57 kg/ha 54 kg/ha	14% 15% - 31% 7%	(2.1.1.2.4) (2.1.1.2.4) - (2.1.1.2.4) (2.1.1.2.4)	52% 53% 17% 59% 51%
Energiforbrug til produktion af N kunstgødning: - ammoniumnitrat mv. - urea - til danske afgrøder - andel af urea til afgrøderne	46 MJ/kg 63 MJ/kg 38 MJ/kg Se tabel 5.2.16	- - - -	- - - -	16% 11% 10%
Energiforbrug til produktion af P kunstgødning	28 MJ/kg	-	-	15%
Energiforbrug til produktion af K kunstgødning	5 MJ/kg	-	(4.2.1.3.3)	29%
Forbrug af kalk pr. ha kalktrængende afgrøde	790 kg/ha	-	(2.2.1.2.1)	10%
Energiforbrug til fremstilling af kalk	0.8 MJ/kg	-	(5.5.5.5.3)	104%
Forbrug af husdyrgødning pr. ha pr. afgrøde beregnes ud fra samme parametre som kunstgødning (se tabel 5.2.10) samt: - total mængde husdyrgødning - mængde N heri	45800 x 10 ⁶ kg 246 x 10 ⁶ kg	- -	(4.1.1.1.1) (4.1.1.1.1)	10% 10%

Data	Middelværdi (μ)	Oprindelig CV	Datakvali- tetsindeks	Korrigeret CV: (og evt. μ)
Energiforbrug til fremstilling af husdyrgødning	35 kWh pr. ton	-	-	Beregnes i afsnit 6.2
Energiforbrug pr. ha til markarbejde for danske afgrøder: - rug - sukkerroe - hvede	2309 MJ/ha 4186 MJ/ha 2836 MJ/ha	20%	(2.1.2.2.1)	23% (og μ 10% mindre)
Energiforbrug til markarbejde pr. ha for udenlandske afgrøder	Se tabel 5.2.23	20%	(3.1.2.5.5)	52% (og μ 10% mindre)
Energiforbug til tørring af korn-afgrøder.	0.24 MJ/kg	-	(4.1.2.1.1)	25% (og μ 10% mindre)

De korrigerede variationskoefficienter, der medregner usikkerheden fra datakvaliteten, varierer fra under 1% (for et datakvalitetsindeks på 1,1,1,1,1) til 104% (i forbindelse med datakvalitetsindekset 5,5,5,5,3). Det skal dog understreges, at der ikke er nogen direkte sammenhæng mellem datakvalitetsindekset og usikkerheden. Således kan to forskellige data med det samme kvalitetsindeks godt have forskellige korrigerede variationskoefficienter.

Det fremgår af nærværende case, at de korrigerede variationskoefficienter domineres af enkelte store usikkerheder. Dette er en naturlig følge af de benyttede forudsætninger, nemlig at de indgående datakvalitetsindikatorer er uafhængige og at alle usikkerhederne antages at være normalfordelt. Et typisk eksempel er energiforbruget til produktion af pesticider, der har en oprindelig variationskoefficient på 58,7%. Når datakvaliteten tages med i betragtning (se tabel 5.2.8), tillægges variationskoefficienterne for de enkelte indikatorer (16,5% og 10%). Alligevel bliver den samlede variationskoefficient kun 62% (kvadratroden af summen af de tre variationskoefficients kvadrater), dvs. en afvigelse på ca. 3% fra den værdi, man ville have fået ved kun at benytte den største usikkerhed (58,7%). I nærværende case er der kun få tilfælde, hvor der er flere store usikkerheder til den samme middelværdi. Et eksempel er forbruget af pesticider til udenlandske afgrøder, hvor både 4. og 5. datakvalitetsindikator får scoren 5 og estimeres til en variationskoefficient på 50% hver (tabel 5.2.7). Dette giver en samlet variationskoefficient på 71%.

Ovenstående observation retsfærdiggør en "hurtig" metode til estimation af de samlede usikkerheder, idet mindre usikkerheder kan ignoreres. Herved vil datakvalitetsindeksene først og fremmest komme til at fungere som "check-lister" til identifikation af de væsentligste usikkerheder for de enkelte data. Herefter kunne der regnes videre med disse væsentlige usikkerheder. Derved ville usikkerheden på det samlede resultat naturligvis blive undervurderet, hvilket kunne imødegås ved introduktion af en erfaringsmæssig baseret korrektionsfaktor. Dette forslag ligner et tilsvarende forslag til overslagsmæssig beregning af usikkerhederne på allerede publicerede livscyklusvurderinger, som fremkom på projektets internationale workshop (se kapitel 7).

5.3 Case 2: Emissioner fra bageri

5.3.1 Beskrivelse af casen

Formålet med nærværende case er at illustrere usikkerheder i industrielle data såvel som i emissioner fra industriel produktion. Som eksempel på industrielle data er produktion af rugbrød valgt. Casen er karakteriseret ved, at der foreligger serier af måledata fra den specifikke virksomhed. Casen tager udgangspunkt i produktion af et konkret produkt - et skiveskåret rugbrød fra Schulstad Brød A/S. Data er taget fra en publiceret livscyklus-screening af dette produkt (Weidema *et al.*, 1995) samt tilgængeligt baggrundsmateriale.

Den funktionelle enhed for produktionen er *1 kg skiveskåret rugbrød*, og i denne case behandles beregningerne af emissionen af BOD og total-N opgjort som g BOD/kg rugbrød og g total-N/kg rugbrød. I beregningerne indgår oplysninger om den årlige produktion af rugbrød, det årlige forbrug af vand, tilsetning af vand ved produktion af rugbrød og emissionen af BOD og total-N pr. liter spildevand.

I tabel 5.3.1 er givet en oversigt over data, der indgår i casen, og umiddelbare kommentarer er knyttet til de enkelte data. Kommentarerne er nærmere uddybet i forbindelse med beskrivelsen af datakvaliteten for de enkelte data.

Tabel 5.3.1

Oversigt over data anvendt i forbindelse med beregning af emissionen af BOD og total-N pr. kg rugbrød.

Data	Kommentarer
Produktion af rugbrød opgjort pr. måned gennem ét år.	Produktionsdata for den totale produktion af rugbrød på Schulstad Brød A/S er fra 1992.
Vandsforbrug for bageriet. Forbruget er opgivet pr. uge gennem 12 måneder (én uge pr. måned).	Fordelingen af vandsforbruget til processer (bageriet) og til sanitære formål mv. er ikke opgivet. Data for vandsforbrug er fra 1992.
Tilsat vandmængde til rugbrød.	Fremgår af opskrift. 33% er angivet som den vandmængde, der tilsettes rugbrød - under bagingen sker der sandsynligvis en fordampning af vand, hvilket dog ikke får betydning for beregningen af spildevandsmængden, da den tilsatte vandmængde forlader fabrikken som damp eller med brødet.
Emission af BOD. 21(20) målinger gennem 14 måneder.	Målingerne af BOD (1991-92) er foretaget på to forskellige udløb, hvori der er målt væsentlig forskellige koncentrationer; vandmængderne i og forskellen mellem de to udledninger er ikke kendt.
Emission af total-N. 13 målinger gennem 12 måneder.	Målingerne af total-N (1991-92) er foretaget på to forskellige udløb, hvori der er målt væsentlig forskellige koncentrationer; vandmængderne i og forskellen mellem de to udledninger er ikke kendt.

5.3.2 Beskrivelse af datakvalitet

Produktion af rugbrød

Produktionen af rugbrød på Schulstad Brød A/S er angivet i ton/måned gennem 12 måneder. Den årlige produktion kan beregnes som $\Sigma p_{\text{måned}} = 21.738$, hvor $p_{\text{måned}}$ er produktionen pr. måned. Den gennemsnitlige produktion pr. måned beregnes som $\Sigma p_{\text{måned}}/12 = 1.812$ ton/måned (svarende til 30,5 dage/måned). Variationen beregnes på månedlige produktionstal korrigert til en gennemsnitsmåned på 30,5 dage. De anvendte data er vist i tabel 5.3.2.

Tabel 5.3.2

Produktion af rugbrød på Schulstad Brød A/S, 1992 angivet i 1.000 kg/måned.

Måned	Total produktion af rugbrød	
	1.000 kg/måned	1.000 kg/måned (30,5 dage)
Januar	1.838	1.808
Februar	1.723	1.812
Marts	1.681	1.654
April	1.990	2.023
Maj	1.683	1.656
Juni	1.700	1.728
Juli	1.766	1.738
August	1.829	1.799
September	1.847	1.878
Oktobre	2.229	2.193
November	1.830	1.860
December	1.622	1.596
<i>Statistisk behandling af data</i>		
antal data	12	12
sum	21.738	21.745
middelværdi	1811,5	1812,1
varians	27.324,27	27.728,63
standardafvigelse	165,30	166,52
variationskoefficient (%)	9,125	9,189

Kvaliteten af data vedrørende den producerede mængde rugbrød kan angives som følger: *Produktion af rugbrød: 1.812 ton/måned (CV: 9,2%) (1,1,1,1,1)*. Begrundelsen for datakvalitetsindex gives i tabel 5.3.3.

Tabel 5.3.3

Datakvalitetsindex for produktion af rugbrød.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed			CV: 9,2%
Pålidelighed	Målte tal. oplysninger leveret af producenten	1	
Fuldständighed	Data over et år vurderes at dække eventuelle sæsonvariationer i forbruget af rugbrød	1	
Tidsmæssig korrelation	Produktionstallene dækker 1992	1	
Geografisk korrelation	Data stammer fra den aktuelle virksomhed	1	
Teknologisk korrelation	Data stammer fra den aktuelle virksomhed. Produktionstallene angiver den samlede produktion af rugbrød på virksomheden.	1	
Samlet usikkerhed			CV: 9,2%

Vandforbrug på bageriet

Det totale vandforbrug er fra virksomheden angivet som m³/uge for 1 uge pr. måned gennem 12 måneder. Det gennemsnitlige vandforbrug beregnes som $\Sigma m_{vand,uge}/52 = 613,7 \text{ m}^3/\text{uge}$ (ton/uge), hvor $m_{vand,uge}$ er vandforbruget pr. uge (s. 127). Det årlige vandforbrug beregnes som $613,7 \text{ ton/uge} \times 52 \text{ uger} = 31.910 \text{ ton/år}$. De anvendte data er vist i tabel 5.3.4.

Tabel 5.3.4

Forbrug af vand på Schulstad Brød A/S, 1992, angivet som m³/uge.

Måned	Forbrug af vand
Januar	735
Februar	673
Marts	556
April	539
Maj	741
Juni	653
Juli	491
August	701
September	647
Oktober	616
November	506
December	506
<i>Statistisk behandling af data</i>	
antal data	12
sum	7364
middelværdi	613,67
varians	8352,6
standardafvigelse	91,39
variationskoefficient (%)	14,9

Kvaliteten af data vedrørende det totale vandforbrug på Schulstad Brød A/S kan angives som følger: *Vandforbrug: 613,7 ton/uge (CV: 15%) (1,1,1,1,1)*. Datakvalitetsindex er vist i tabel 5.3.5.

Tabel 5.3.5

Datakvalitetsindex for det totale vandforbrug på Schulstad Brød A/S.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed			CV: 15%
Pålidelighed	Oplysninger leveret af producenten	1	
Fuldstændighed	12 stikprøver (af en uges varighed) udtaget jævnligt over hele året vurderes at udgøre en repræsentativ stikprøve (23% af den samlede population er undersøgt)	1	
Tidsmæssig korrelation	Vandforbrugstallene dækker 1992	1	
Geografisk korrelation	Data stammer fra den aktuelle virksomhed	1	
Teknologisk korrelation	Data stammer fra den aktuelle virksomhed. Vandforbruget angiver det samlede vandforbrug til både produktion som til sanitære formål	1	
Samlet usikkerhed			CV: 15%

Tilsat vandmængde til rugbrød

Den mængde vand, der tilsættes rugbrødsdejen, er oplyst af Schulstad Brød A/S (jf. opskrift). Variationen i doseringen af vand ved produktion af rugbrød er ikke oplyst, men den antages at være 1%. Kvaliteten af data vedrørende forbrug af vand i rugbrød kan angives, som følger: *Vandforbrug til rugbrød: 0,33 kg vand/kg brød (CV: 1%; estimeret) (1,1,1,1,1)*. Datakvalitetsindex er vist i tabel 5.3.6.

Tabel 5.3.6

Datakvalitetsindex for tilsat vandmængde til rugbrød.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed			CV: 1%, estimeret
Pålidelighed	Oplysninger leveret af producenten (jf. opskrift)	1	
Fuldständighed	Opskrifter udvikles/ændres løbende, men den angivne opskrift må formodes at være repræsentativ for det betragtede tidspunkt	1	
Tidsmæssig korrelation	Opskriften er den aktuelle for 1992	1	
Geografisk korrelation	Data stammer fra den aktuelle virksomhed	1	
Teknologisk korrelation	Data stammer fra den aktuelle virksomhed	1	
Samlet usikkerhed			CV: 1%

Udledningen af BOD og total-N

Udledningen af BOD og total-N er målt i perioden juni 1991 - august 1992 henholdsvis september 1991 - august 1992 i to udledninger.

Vandmængderne i og forskellen mellem de to udledninger er ikke kendt, hvorfor fordelingen af vandmængderne mellem de to udledninger antages at være 1:1. Beregning af den gennemsnitlige udledning af BOD er udført som følger: $(\Sigma \text{BOD}_N/n_N + \Sigma \text{BOD}_\theta/n_\theta)/2 = (46.240/21 + 1.569/20)/2 = 1.140 \text{ mg BOD/l}$, hvor BOD_N og BOD_θ er indholdet af BOD i udledningen til henholdsvis nord og øst; n angiver antallet af målinger på de pågældende udløb. Beregning af den gennemsnitlige udledning af total-N er udført som følger: $(\Sigma \text{total-N}_N/n_N + \Sigma \text{total-N}_\theta/n_\theta)/2 = (862/13 + 61/13)/2 = 35,5 \text{ mg total-N/l}$, hvor total-N_N og total-N_θ er indholdet af total-N i udledningen til henholdsvis nord og øst. Middelværdi, standardafvigelse og variationskoefficient er beregnet for koncentrationen af BOD og total-N i hvert enkelt udløb; herefter er middelværdien af koncentrationen af BOD og total-N beregnet for de to udløb, og den samlede standardafvigelse og variationskoefficient er beregnet ud fra standardafvigelserne på de enkelte udløb. De anvendte data og beregninger er vist i tabel 5.3.7.

Tabel 5.3.7

Emissioner af BOD og total-N fra Schulstad Brød A/S (udløb nord og øst), 1991 - 92 angivet i mg/l.

Dato	BOD (mg/l)		total-N (mg/l)	
	nord	øst	nord	øst
92.08.06	3.000	6	75.6	4.5
92.06.11	2.000	100	-	-
92.06.04	890	88	31.2	12.2
92.05.06	980	19	3.2	0.1
92.04.23	870	-	-	-
92.04.02	1.500	17	4.7	0.8
92.03.19	2.100	110	55	0.1
92.02.04	3.700	110	112.3	1.6
92.01.23	800	270	-	-
92.01.15	1.200	33	16.7	0.1
91.12.04	2.200	140	-	-
91.11.21	1.300	78	47.6	8.4
91.11.07	4.600	38	257	0.1
91.10.23	2.100	200	61	27
91.10.08	3.400	110	81.3	3.2
91.09.25	2.100	6	68.6	0.8
91.09.11	1.900	9	48	2.1
91.08.22	5.800	8	-	-
91.07.25	1.300	79	-	-
91.07.10	1.200	38	-	-
91.06.27	3.300	110	-	-
<i>Statistisk behandling af data for hvert enkelt udløb</i>				
antal data	21	20	13	13
sum	46.240	1.569	862	61
middelværdi	2.202	78.5	66.3	4.7
varians	1690.000	4.638	3.931	53.29
standardafvigelse	1.300	68.1	62.7	7.3
variationskoefficient (%)	59	87	95	155
<i>Statistisk behandling af data for begge udløb samlet</i>				
antal data	2	2	2	2
sum	2.280.5	2.280.5	71	71
middelværdi	1.140	1.140	35.5	35.5
varians	423.659	423.659	996.1	996.1
standardafvigelse	650.9	650.9	31.6	31.6
variationskoefficient (%)	57	57	89	89

- Data mangler.

Kvaliteten af data vedrørende emissioner af BOD og total-N fra Schulstad Brød A/S kan angives som følger: *Udledning af BOD: 1,14 g/l (CV: 57%) (1,1,1,2,1)* og *udledning af total-N: 0,0355 g/l (CV: 89%) (1,1,1,2,1)*. Datakvalitetsindex er vist i tabel 5.3.8.

Den supplerende usikkerhed som følge af antagelsen om ligelig fordeling af spildevand mellem udløb nord og øst, er estimeret ud fra en "worst case" betragtning. I værste tilfælde antages 90% af spildevandet at indeholde den høje koncentration af BOD og total-N. Middelværdierne for worst case betragtningen kan estimeres til 1.990 mg BOD/l og 60,1 mg total-N/l. Disse værdier antages at udgøre øverste grænse i 95% konfidensintervallet for middelværdien for 1:1 fordelingen. Ud fra denne antagelse kan de supplerende variationskoefficienter estimeres til henholdsvis 37% og 35%.

Tabel 5.3.8
Datakvalitetsindex for emission af BOD og total-N.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed <i>BOD</i> <i>total-N</i>			CV: 57% CV: 89%
Pålidelighed	Målte tal, oplysninger leveret af producenten	1	
Fuldstændighed	Målinger er foretaget gennem 12 - 14 måneder, hvorfor de vurderes at dække over evt. sæsonvariationer.	1	
Tidsmæssig korrelation	Data stammer fra perioden juni 91 - august 92 - selv om perioden ikke er sammenfaldende med perioden for produktionsoplysningerne vurderes de foreliggende data at være repræsentative.	1	
Geografisk korrelation <i>BOD</i> <i>total-N</i>	Data stammer fra den aktuelle virksomhed. Målingerne af BOD og total-N er foretaget på to forskellige udledninger, hvori der er målt væsentlig forskellige koncentrationer; fordelingen af vandmængderne mellem de to udledninger er antaget at være 1:1	2	CV: 37% ¹ CV: 35% ¹
Teknologisk korrelation	Data stammer fra den aktuelle virksomhed.	1	
Samlet usikkerhed <i>BOD</i> <i>total-N</i>			CV: 68% CV: 96%

1. Usikkerheden er estimeret ud fra en "worst case" betragtning om, at 90% af spildevandsmængden udlødes via det nordlige udløb.

Tages højde for supplerende usikkerheder, kan emissionerne af BOD og total-N angives som følger: *Udledning af BOD: 1,14 g/l (CV: 68%) (1,1,1,m,1)* og *udledning af total-N: 0,0355 g/l (CV: 96%) (1,1,1,m,1)*.

5.3.3 Samlet vurdering af datakvaliteten

Datakvaliteten for de indgående parametre er sammenfattet i tabel 5.3.9. Datakvaliteten er beskrevet i forhold til den funktionelle enhed, som er 1 kg brød i den oprindelige undersøgelse (Weidema *et al.*, 1995). I principippet kan datakvaliteten beskrives ud fra årlig produktion som funktionel enhed. Valg af årlig produktion som funktionel enhed vil ikke ændre på datakvaliteten, men en mellemregning vil kunne udelades.

I den gennemgåede case er datakvaliteten af alle de indgående data beskrevet.

Tabel 5.3.9
Middelværdi, variationskoefficient (CV), datakvalitetsindex og total variationskoefficient for de indgående data.

Parameter	Middelværdi	Oprindelig usikkerhed	Datakvalitetsindex	Samlet usikkerhed (og evt. korrigert middelværdi)
Produktionsmængde	1.812 ton/måned	9,2%	1.1.1.1.1	9,2%
Vandforbrug	613,7 ton/uge	15%	1.1.1.1.1	15%
Vand tilsat ved produktion af rugbrød	0,33 kg vand/kg brød	1%, estimeret	1.1.1.1.1	1%
Udledning af BOD	1,14 g/l	57%	1.1.1.2.1	68%
Udledning af total-N	0,036 g/l	89%	1.1.1.2.1	96%
Emissionsfaktorer for BOD og total-N	1,30 g BOD/kg brød og 0,040 g total-N/kg brød			Beregnet i afsnit 6.3

I de tilfælde, hvor datakvaliteten er god, kan datakvaliteten evt. gives for aggregerede data, som vist i tabel 5.3.10.

Tabel 5.3.10
Middelværdi, variationskoefficient (CV), datakvalitetsindex og total variationskoefficient for aggregerede data.

Parameter	Middelværdi	Oprindelig usikkerhed	Datakvalitetsindex	Samlet usikkerhed (og evt. korrigert middelværdi)
Spildevand pr. produceret mængde rugbrød	1,14 kg spv./kg brød		1.1.1.1.1	
Udledning af BOD	1,14 g/l	57%	1.1.1.2.1	68%
Udledning af total-N	0,036 g/l	89%	1.1.1.2.1	96%
Emissionsfaktorer for BOD og total-N	1,30 g BOD/kg brød og 0,040 g total-N/kg brød			Beregnet i afsnit 6.3

I det viste eksempel kan datakvalitetsindexet umiddelbart videreføres fra de oprindelige til de aggregerede data, omend bestemmelsen af variationskoefficienten for de aggregerede data kræver statistiske beregninger. I de tilfælde, hvor de indgående data har en dårlig datakvalitet, risikerer man at sløre oprindelsen til den dårlige datakvalitet, hvis datakvaliteten kun angives for de aggregerede data.

I tilfælde, hvor datakvaliteten er ringe, kræves der en mere detaljeret beskrivelse af datakvaliteten/anvendelse af datakvalitetsindikatorerne. I de tilfælde vil et af formålene med anvendelse af datakvalitetsindika-

torerne være at bruge dem til at spore oprindelsen til den dårlige data-kvalitet og have mulighed for at forbedre datakvaliteten ved fornyet dataindsamling.

Datakvalitetsindekset er ikke i sig selv tilstrækkelig dokumentation for usikkerhed, men kun en indikation af hvilke data, der er behæftet med usikkerhed. Dvs. den sammenfattende tabel erstatter ikke den mere detaljerede beskrivelse af data.

5.4 Case 3: Energiforbrug til nedkøling af varer

5.4.1 Beskrivelse af casen

Nærværende case er en illustration af betydningen af levetiden af køleskabe på estimation af energiforbruget til nedkøling af varer i private husholdninger. Casen tager udgangspunkt i energiovervejelser gennemført i forbindelse med livscykluusscreening af udvalgte fødevarer. Data er taget fra Weidema *et al.* (1995) samt fra tilgængeligt baggrundsmateriale.

Energiforbruget for køleskabe er valgt som eksempel, idet det giver mulighed for overvejelser vedrørende forskellige fordelinger af levetider for husholdningsapparater samt overvejelser om allokering. I beregningerne indgår oplysninger om salg af køleskabe i Danmark i perioden 1975 - 91, energiforbrug for et gennemsnitskøleskab samt en opgørelse over energiforbrug for ældre køleskabe. Eksemplet giver ligeledes mulighed for at belyse problematikken omkring allokering, idet energiforbruget for køleskabe afhænger af en lang række faktorer:

- energiforbrugets afhængighed af frekvensen for supplering/udskiftning af varer i køleskabet herunder energiforbrug til nedkøling af varen
- energiforbrug til at holde varen nedkølet til ca. 5°C
- energiforbrug i forbindelse med døråbninger
- energiforbrugets afhængighed af fyldningsgraden af køleskabet

Køleskabet afgiver endvidere varme, som i principippet fortrænger energiforbrug til rumopvarmning. Den fortrængte energi skal dersør godskrives energiforbruget til af holde de aktuelle varer nedkølede.

Den funktionelle enhed i den oprindelige livscykluusscreening er husholdsvist 1 kg brød og 1 kg skinke. Spørgsmålet omkring allokering af energiforbruget kan herefter præsiceres til

hvor stor en del af energiforbruget til drift af et køleskab skal allokeres til at holde 1 kg skinke nedkølet til den ønskede temperatur?

I tabel 5.4.1 er givet en oversigt over de data, der indgår i casen, og umiddelbare kommentarer er knyttet til de enkelte data. Kommentarerne er nærmere uddybet i forbindelse med beskrivelsen af datakvaliteten for de enkelte data.

Salgsstatistikken for køleskabe (1975 - 91) er anvendt som udgangspunkt for estimation af aldersfordelingen af køleskabe i funktion i 1992 ud fra forudsætningen om et mættet marked. Energiforbruget for køleskabe af forskellig alder, som er i funktion i 1992, er estimeret ud fra kendskab til energiforbruget for et nyt 200 l køleskab (1991), variation i energiforbruget for forskellige køleskabe af samme model (1987 - 89) samt energiforbrug for nye køleskabe af forskellig alder. I den aktuelle case er det godtgjort, at det er tilstrækkeligt at allokerere energiforbrug til nedkøling af 1 kg skinke i forhold til det rumfang produkterne optager i køleskabet. Herved anvendes et energiforbrug udtrykt som kWh/l/dag, hvori der indgår rumfanget af varen i liter i forhold til køleskabets samlede rumfang og antal dage varen opbevares under køl.

Den valgte allokering er underbygget med en række overslagsmæssige betragtninger ("worst case").

Til underbyggelse af den valgte allokering indgår oplysninger om energiforbrug pr. døråbning, antal døråbninger pr. medlem i husstanden, gennemsnitlig varmekapacitet for fødevarer, udnyttelsesgrad for køleskabe samt fortrængt energi til rumopvarmning.

Tabel 5.4.1

Oversigt over data anvendt i forbindelse med beregning af energiforbruget til nedkøling af varer i private husholdninger.

Data	Kommentarer
Salgsstatistik for køleskabe, 1975-92.	Perioden 1975-91 dækker rimeligt den periode, hvor køleskabe, der fungerer i 1992, kan forventes at være købt.
Energiforbrug for et gennemsnitskøleskab.	Energiforbruget for et 200 liters køleskab er estimeret ud fra en gennemsnitlig model på markedet i 1991 ved hjælp af et simuleringssprogram udviklet på Fysisk Lab. III, DTU.
Variation på energiforbruget for køleskabe af én model.	Et lavenergikøleskab er i 1987-89 testet hos 116 husstande af forskellig størrelse og bl.a. energiforbruget er bestemt. Variationen antages at kunne bruges som udtryk for generel variation på energiforbruget for køleskabe af en given type.
Energiforbrug for ældre køleskabe.	Ud fra data i en schweizisk database er energiforbruget for ældre køleskabe estimeret.
Energiforbrug til døråbning.	I forbindelse med testen af lavenergikøleskabet blev antallet af døråbninger/døgn, døråbningsstid/døgn og energiforbrug/år undersøgt. Her ud fra blev energiforbruget estimeret pr. sekund døren var åben.
Varmekapacitet for fødeværer (gennemsnit).	Den gennemsnitlige varmekapacitet for fødeværer er estimeret af forfatterne.
Udnyttelsesgrad for køleskab.	I forbindelse med testen af lavenergikøleskabet blev 114 husstande (med lavenergikøleskab) spurgt om udnyttelsesgraden for deres køleskab.
Fortrængt energi til rumopvarmning.	Oplysninger om fortrængt energi til rumopvarmning ved anvendelse af forskellige former for køleskabe er baseret på indhentede oplysninger (kvalificeret estimat) og derfor ikke nærmere underbyggede.
Vægtfylde for rugbrød og skinke.	Vægtfonden for skinke er målt af producenten.
Opbevaringstid for skinke.	Opbevaringstiden for skinke er skønnet af forfatterne.

5.4.2 Beskrivelse af datakvalitet

Salgsstatistik for køleskabe

Salgstal for køleskabe dækende perioden 1975-91 er oplyst af FEHA (Foreningen af Fabrikanter og Importører af Elektriske Husholdningsapparater). Salgstallene er angivet i tabel 5.4.2.

Tabel 5.4.2

Salgsstatistik for køleskabe. Det estimerede energiforbrug for ældre køleskabe samt andelen af tilbageværende køleskabe af forskellig drøgång i 1992 er ligeledes vist.

Salgsår	Antal solgt (1.000 stk)	Energiforbrug (kWh/år/køleskab)	Levetiden antaget at følge en normalfordeling ²		
			% i funktion i 1992	Antal i funk- tion i 1992 (1.000 stk)	Energiforbrug MWh
1991	194	277,0	100	194,00	53.738,00
1990	186	287,5	100	186,00	53.475,00
1989	186	298,0	100	186,00	55.428,00
1988	152	308,5	100	152,00	46.892,00
1987	162	319,0	100	162,00	51.678,00
1986	175	329,5	99	173,25	57.085,88
1985	161	340,0	98	157,78	53.645,20
1984	167	350,5	96	160,32	56.192,16
1983	150	361,0	92	138,00	49.818,00
1982	143	371,5	87	124,41	46.218,32
1981	145	382,0	81	117,45	44.865,90
1980	152	392,5	72	109,44	42.955,20
1979	174	403,0	61	106,14	42.774,42
1978	166	413,5	50	83,00	34.320,50
1977	160	424,0	39	62,40	26.457,60
1976	175	434,5	28	49,00	21.290,50
1975	155	445,0	19	29,45	13.105,25
1974	155 ¹	455,5	13	20,15	9.178,33
1973	155 ¹	466,0	7	10,85	5.056,10
1972	155 ¹	476,5	4	6,20	2.954,30
1971	155 ¹	487,0	2	3,10	1.509,70
1970	155 ¹	497,5	1	1,55	771,13
Total				2226	769.409,49

1. Statistikken går tilbage til 1975: salgstal for tidligere år er skønnet til samme niveau.
2. Normalfordelingen er antaget at have en middelværdi på 14 og en varianskoeficient på 25% dvs. $X \in (14,3,5^2)$.

Datakvalitetsindex for salgstal for køleskabe kan angives som $(2, 1, 1, 1, 1)$. Begrundelsen for det givne datakvalitetsindex er vist i tabel 5.4.3. Usikkerheden på salgstallene er ikke angivet i kilden, men den antages at være 1% (CV) på de årlige statistiske opgørelser pga. fejl i indberetninger mv.

Tabel 5.4.3
Datakvalitetsindex for salgstal for køleskabe.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed			
Pålitelighed	Indberetningerne til statistikken er ikke verificeret.	2	CV: 1%. estimeret
Fuldstændighed	Der foreligger tal fra alle år undtaget årene 1970-74 (som ud fra den antagne fordeling - se senere - dækker 2% af de fungerende køleskabe i 1992).	1	
Tidsmæssig korrelation	Statistikken dækker det tidsrum, hvor et køleskab, der fungerer i 1992, kan forventes at være købt undtaget årene 1970-74. På trods af, at 2% af de fungerende køleskabe ikke er omfattet, vurderes statistikken at være dækkende.	1	
Geografisk korrelation	Salgstallene dækker det danske marked.	1	
Teknologisk korrelation	Idet salgstallene dækker perioden 1975-91 må de siges at dække det aktuelle teknologiske niveau. En eventuel fejl kan introduceres ved fastlæggelse af energiforbruget for ældre køleskabe år 1992 - se senere.	1	
Samlet usikkerhed			CV: 1%

Tages højde for den supplerende usikkerhed kan datakvaliteten for salgstallene angives som (CV: 1%) (m, 1, 1, 1, 1).

Aldersfordeling af køleskabe i brug i 1992

Ud fra antagelsen om et mættet marked, er antallet af køleskabe i funktion antaget at være det samme som antallet af husstande i Danmark. I løbet af 1991 og de foregående 13 år er der solgt et antal køleskabe svarende til det samlede antal køleskabe, hvorfor gennemsnitslevetiden for køleskabe antages at være 14 år. Nogle køleskabe vil holde i længere tid, mens andre vil holde i kortere tid. Levetiden er antaget at følge en normalfordeling med gennemsnitslevetid på 14 år.

Der vil være en usikkerhed på antagelsen om et mættet marked, idet nogle husstande vil have to køleskabe. Hvis 2% af de årligt indkøbte køleskabe svarende til 4.000 køleskabe ikke erstatter et gammelt køleskab, vil det give en forøgelse af gennemsnitsalderen på 0,3 måned dvs. være uden nævneværdig betydning for gennemsnitsalderen. Det må ydermere påregnes, at der står en del køleskabe i fritidshuse (som ikke indgår i den anvendte opgørelse over husstande i Danmark), og det drejer sig om nye såvel som ældre køleskabe forstået på den måde, at det "gamle" køleskab ofte flyttes til fritidshuset ved anskaffelse af et nyt til helårbeboelsen. Forholdet mellem nye og gamle køleskabe i fritidshuse kendes ikke, hvorfor der med den foreliggende viden ikke kan siges noget om betydningen af køleskabe placeret i fritidshuse for aldersfordelingen af køleskabe placeret i helårshuse.

Som et alternativ antages, at levetiden følger en fordeling med gennemsnitslevetiden 14 år; i dette tilfælde kan medianlevetiden være forskellig fra gennemsnitslevetiden. I afsnit 6.4.1 sammenlignes normalfordelingen med en gammafordeling. Antagelsen om fordelingen

af levetiden omkring gennemsnitslevetiden (normalsfordeling, gamma-fordeling eller en tredje fordeling) kan ikke tildeles nogen usikkerhed som sådan, idet antagelsen enten er rigtig eller forkert. Det kan ikke med de foreliggende data verificeres hvilken fordeling, der er den sande.

Energiforbrug for et nyt køleskab årgang 1991

Energiforbruget er estimeret for et gennemsnitligt køleskab årgang 1991 ved hjælp af et simuleringsprogram (Pedersen, 1992). Simuleringsprogrammet er udviklet af Laboratoriet for Køleteknik, DTU, og simuleringerne er udført af Fysisk Laboratorium III. Udgangspunkt for simuleringer af energiforbruget under forskellige forhold er oplysninger om køleskabe på de danske marked. Det er ikke oplyst hvilke mærker, modeller mv., der indgår i estimationen af gennemsnitsenergiforbruget. 64 forskellige modeller (90 - 314 liter) indgår i opstillingen af en regressionslinie for energiforbruget:

$$E(vol) = 0,52 + 0,0012 \times V [kWh/24 timer] \quad (17)$$

Heraf kan energiforbruget for drift af et køleskab på 200 liter estimeres til 0,76kWh/24 timer. De anvendte data til opstilling af regressionslinien samt statistiske data er vist i tabel 5.4.4.

Tabel 5.4.4

Data anvendt til opstilling af regressionslinie for det gennemsnitlige energiforbrug for køleskabe (uden frostbox) på det danske marked årgang 1991 samt oplysninger om den statistiske analyse af de foreliggende data (Pedersen, 1992). Den statistiske behandling af data er foretaget på grundlag af de opgivne værdier.

Vol.	kWh/dag	Vol.	kWh/dag	Vol.	kWh/dag	Vol.	kWh/dag	Vol.	kWh/dag
245	0,85	215	0,80	250	0,70	152	0,65	244	0,90
164	0,85	185	0,70	-	-	314	1,10	244	0,90
245	0,85	143	0,65	140	0,80	247	0,85	113	0,65
160	0,70	170	0,80	220	0,70	244	0,85	142	0,70
160	0,75	248	0,95	128	0,61	90	0,70	215	0,80
210	0,63	248	0,95	199	0,57	160	0,65	277	0,80
230	0,90	203	0,80	244	0,60	210	0,80	160	0,75
230	0,90	183	0,70	172	0,70	255	0,90	160	0,75
171	0,75	146	0,75	200	0,28	160	0,70	160	0,75
242	0,85	239	0,80	274	0,75	248	0,75	160	0,80
170	0,75	150	0,70	242	0,80	150	0,75	210	0,62
215	0,80	230	0,85	242	0,80	164	0,70	230	0,79
246	0,85	210	0,63	204	0,60	198	0,90	245	0,81
<i>Statistisk behandling af data</i>									
Afskæring									0,51055
Hældning									0,00124
Standardafvigelse på energiforbrug, s_y									0,11787
Standardafvigelse på hældningskoefficient, s_b									0,00029
Korrelationskoefficient (r)									0,48
Forklарingsgrad (R^2)									0,2317
Gennemsnitligt volumen									201,953
Antal observationer, n									64

Det årlige gennemsnitsenergiforbrug for et køleskab på 200 liter kan estimeres til 277 kWh/år/køleskab. Dette tal er behæftet med usikkerheden fra estimationen af energiforbruget ud fra regressionslinien samt fra variationen i energiforbruget til samme type køleskab placeret hos forskellige forbrugere.

Usikkerheden på de estimerede energiforbrug for det givne køleskab på 200 l er estimeret ved

$$s(\hat{y} \mid X^* = 200^\circ C) = s_y \sqrt{\frac{1}{n} + s_b \sqrt{(x^* - \bar{x})^2}} \quad (18)$$

hvor s_y og s_b er henholdsvis standardafvigelse for energiforbruget og hældningskoefficienten
 n er antal observationer
 \bar{x} er gennemsnitsvolumen for køleskabene

Usikkerheden for det estimerede energiforbrug for et køleskab på 200 liter er 4,756 kWh/år/køleskab. Heraf kan variationskoefficienten for det estimerede energiforbrug estimeres til 1,7%.

Variationen i energiforbruget for et gennemsnitskøleskab på 200 l placeret hos forskellige forbrugere kendes ikke. Med henblik på at estimere denne variation tages udgangspunkt i en undersøgelse, hvor energiforbruget for én type køleskab er undersøgt hos en række forbrugere. I forbindelse med test af et lavenergikøleskab hos 116 husstande af forskellig størrelse er det årlige energiforbrug bestemt (Pedersen & Lawætz, 1989). Variationskoefficienten for disse 116 ens køleskabes energiforbrug anvendes som estimat for den tilsvarende variation på energiforbruget for gennemsnitskøleskabet. De statistiske data for lavenergikøleskabet er vist i tabel 5.4.5.

Tabel 5.4.5
Statistiske data for det årlige energiforbrug for et lavenergikøleskab testet hos 116 husstande.

antal observationer	116
sum	10379,8
middelværdi	89,48
varians	239,125
standardafvigelse	15,464
variationskoefficient (%)	17

Variansen på regressionslinien og variansen på energiforbruget på et køleskab antages at være uafhængige af hinanden og dermed additive, hvorfor den samlede variationskoefficient for energiforbruget for et køleskab kan udtrykkes som $((1,7\%)^2 + (17\%)^2)^{1/2} = 17,1\%$.

Kvaliteten af data vedrørende det gennemsnitlige energiforbrug for et køleskab årgang 1991 kan angives som følger: 277 kWh (CV: 17%) (2,1,1,1,1). Datakvalitetsindex er vist i tabel 5.4.6. Energiforbruget for de enkelte modeller er oplyst af producenten og ikke verificeret af en uafhængig instans, hvorfor energiforbruget antages at være behæftet med en supplerende usikkerhed på 5% (CV).

Tabel 5.4.6
Datakvalitetsmatrix for energiforbrug for køleskabe årgang 1991.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed ¹			CV: 17%
Pålitelighed	Energiforbruget for de enkelte modeller er oplyst af producenterne men er ikke verificeret af uafhængig instans	2	CV: 5%
Fuldstændighed	65 modeller (af forskellige mærker) indgår i opstillingen af regressionslinien for energiforbruget	1	
Tidsmæssig korrelation	De 65 modeller var på de danske marked i juli 1991	1	
Geografisk korrelation	ditto	1	
Teknologisk korrelation	Energiforbruget er beregnet på grundlag af oplysninger om 65 modeller på det danske marked i juli 1991. Der kunne muligvis være sket en reduktion i energiforbruget ved nye modeller inden 1992, men hovedparten af modellerne fra 1991 vurderes stadig at være på markedet i 1992.	1	
Samlet usikkerhed			CV: 18%

1. Den "oprindelige usikkerhed" er beregnet ud fra usikkerheden på regressionen, hvorved energiforbruget for et 200 l køleskab er bestemt samt usikkerheden på energiforbruget til én type køleskabe placeret hos en række forskellige forbrugere.

Tages højde for supplerende usikkerheder kan det gennemsnitlige energiforbrug for et køleskab (200 l) årgang 1991 angives som: 277 kWh (CV: 18%) (m, 1, 1, 1, 1).

Energiforbrug for ældre køleskabe

Der foreligger ingen danske undersøgelser/opgørelser af energiforbruget for ældre køleskabe, hvorfor der tages udgangspunkt i en schweizisk opgørelse. På grundlag af en schweizisk database, som indeholder oplysninger om energiforbruget for nye køleskabe af forskellig alder, er stigningsfaktoren for energiforbruget estimeret (Pedersen et al., 1993). De anvendte data er vist i tabel 5.4.7.

Tabel 5.4.7
Energiforbrug for køleskabe af forskellig alder (Pedersen et al., 1993).

Salgsår	Energiforbrug Kwh/år	Forskel i energiforbrug kWh/år	Fald i energiforbrug kWh/år
1992	317	21	10.5
1990	338	52	10.5
1985	390	53	10.5
1980	443	52	10.5
1975	495	52	10.5
1970	547		

Kvaliteten af data (en konstant) til estimation af energiforbrug for køleskabe af en given årgang kan angives som: +10,5 kWh/dr (CV: 17%) (2,2,1,3,1), idet variationen i det aktuelle energiforbrug for et nyt køleskab (og hermed stigningen for ældre køleskabe som følge af forbedringer) antages at være den samme, som estimeret for det danske lavenergikøleskab. Datakvalitetsindex er vist i tabel 5.4.8.

Effektiviteten af køleskabe formodes at falde med årene, dvs. energiforbruget stiger med alderen. Konstanten vurderes derfor at være behæftet med en usikkerhed på 10% (CV) som følge af manglende fuldstændighed (score: 2). Der foreligger ikke undersøgelser, der viser faldende effektivitet. Opszumningsmidlet i isoleringen kan forventes at være fastholdt i køleskabets levetid, mens køleskabet kan betragtes som defekt fra det øjeblik, hvor kølemidlet begynder at sive fra kølesystemet (Preben Buhl Pedersen, Dansk Energi Analyse A/S, personlig information, september 1995).

Ved beregning af det samlede energiforbrug for ældre køleskabe kan der tages højde for en evt. faldende effektivitet ved at korrigere energiforbruget for et køleskab med en given alder med f.eks. 1% pr. år. I nærværende eksempel er denne korrektion udeladt.

Tabel 5.4.8

Datakvalitetsindex for energiforbrug for ældre køleskabe.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed			CV: 17%, estimeret
Pålitelighed	Der kan ikke siges noget om påliteligheden ud fra det foreliggende materiale, men oplysningerne stammer formodentlig fra producenter.	2	CV: 5%, estimeret
Fuldstændighed	Informationerne dækker hele den relevante periode, men der er sandsynligvis tale om energiforbruget for nye køleskabe, dvs. der er ikke taget højde for at energiforbruget for det enkelte køleskab er stigende med alderen (effektiviteten falder).	2	CV: 10%, estimeret Energiforbrug korrigeres evt. med 1%/år
Tidsmæssig korrelation	Informationerne dækker hele den relevante periode.	1	
Geografisk korrelation	Oprindelsstedet for data er Schweiz.	3	CV: 2,5%, estimeret
Teknologisk korrelation	Informationerne formodes at dække de nyeste modeller på det givne tidspunkt.	1	
Samlet usikkerhed			CV: 21% eller ved kor. af middelværdi: 18%

Når der tages højde for supplerende usikkerheder, kan konstanten anvendt til estimation af energiforbrug for køleskabe af en given alder angives som: + 10,5 kWh/år (CV: 21%) (m,m,1,m,1).

Samlet energiforbrug for køleskabe i 1992

Det årlige energiforbrug for et nyt køleskab af en given alder (T [år]) kan angives som:

$$E(T) = 277 \text{ kWh/år} (\pm 18\%) + T \times 10,5 \text{ kWh/år} (\pm 21\%) \quad (19)$$

Under forudsætning af en faldende effektivitet på 1% pr. år kan energiforbruget i 1992 for et køleskab af en given alder angives som:

$$E(T) = (277 \text{ kWh/år} (\pm 18\%) + (T \times 10,5 \text{ kWh/år} (\pm 18\%))) \times (1 + (T \times 0,01 \text{ kWh/år})) \quad (20)$$

I den videre beregning er denne korrektion ikke anvendt.

Samlet energiforbrug for køleskabe i 1992

Det samlede energiforbrug for køleskabe er beregnet som summen af energiforbruget til køleskabe af forskellig alder (over perioden 1970 - 91); se tabel 5.4.2, hvorfra det fremgår, at det samlede årlige energiforbrug i 1992 er beregnet til 769.410.000 kWh/år. Ved beregning af usikkerheden på det samlede årlige energiforbrug (se afsnit 6.4) medtages usikkerheden på salgsstatistikken, som er estimeret 1%, jf. tabel 5.4.3.

Gennemsnitligt energiforbrug for køleskabe

Det gennemsnitlige energiforbrug for køleskabe beregnes ud fra det samlede årlige energiforbrug for køleskabe i 1992 og antallet af køleskabe i 1992: 769.410.000 kWh/år / 2.226.000 køleskabe = 345,6 kWh/køleskab/år svarende til 0,95 kWh/dag eller under forudsætning af, at køleskabet rummer 200 liter: 4,73 Wh/liter/dag.

Usikkerhed ved allokering af energiforbruget til den enkelte fødevare
 I Weidema et al. (1995) præsenteres en række supplerende data vedrørende energiforbrugets sammenhæng med døråbning, varmekapacitet for fødevarer, udnyttelsesgrad for køleskabe og fortrængt energi til rumopvarmning. Disse data anvendes alle til at begrunde den foretagne allokering af gennemsnits-energiforbruget på 4,73 Wh/liter/døgn direkte i forhold til det rumfang, den aktuelle fødevare optager. Argumentet for denne allokering er, at energiforbruget primært hænger sammen med det rumfang, der skal holdes nedkølet. Det angivne energiforbrug indbefatter imidlertid også energiforbruget betinget af døråbning og nedkøling af varer. Disse forhold hænger ikke sammen med den enkelte vares rumfang, og kunne dermed begrunde en allokering af de respektive andele af energiforbruget i forhold til, hvor tit den enkelte vare tages ind og ud af køleskabet, samt i forhold til den enkelte vares varmekapacitet. Endvidere afhænger allokeringen af udnyttelsesgrad og fortrængt energi til rumopvarmning. Disse forhold vil alle betyde en forøget usikkerhed på det opgivne energiforbrug. I tabel 5.4.9 er de supplerende data sammenfattet. De angivne data og usikkerheder forklares nærmere i det følgende.

Tabel 5.4.9

Oversigt over supplerende data og faktisk anvendte data.

Parameter	Værdi	Usikkerhed (CV)
Gennemsnitligt energiforbrug for køleskabe per liter køleskab	4,73 Wh/liter/døgn	
Energiforbrug ved døråbning	18,8 kWh/køleskab/år	5,4%
Energiforbrug til nedkøling af fødevarer	0,15 kWh/døgn	11%
Udnyttelsesgrad for køleskabe	65%	15%
Fortrængt energi til opvarmning	÷ 50%	21%
Gennemsnitligt energiforbrug for køleskabe allokeret pr. liter fødevarer ¹	3,64 Wh/liter/døgn	

1. Energiforbruget pr. liter fødevarer er korrigert for en udnyttelsesgrad på 65% og en fortrængt mængde energi til rumopvarmning på 50%.

Ved allokering af energiforbrug til køleskabe til energiforbrug pr. liter fødevarer skal energiforbruget pr. liter køleskab korrigeres for udnyttelsesgrad og fortrængt energi til rumopvarmning. Korrigeringen for udnyttelsesgraden foretages som følger: $4,73 \text{ Wh/liter/døgn} / 65\% = 7,28 \text{ Wh/liter fødevarer/døgn}$. 50% af den anvendte energi til køleskabet går til reduktion af energiforbrug til rumopvarmning, hvorför det reelle energiforbrug til nedkøling af fødevarer er $7,28 \text{ Wh/liter fødevarer/døgn} \times 50\% = 3,64 \text{ Wh/liter fødevarer/døgn}$.

Energiforbrug til døråbning

I forbindelse med test af et lavenergikøleskab LER200 med nettovolumenet 200 liter (Pedersen & Lawætz, 1989) blev sammenhængen mellem den daglige døråbningstid og det årlige energiforbrug bestemt, og heraf blev følgende parametre bestemt:

- antal døråbninger/døgn (middelværdi: 27 døråbninger/døgn)
- samlet døråbningstid/døgn (middelværdi: 261 sekunder/døgn)
- energiforbrug/år (middelværdi: 88,8 kWh/år)

Ud fra disse parametre blev energiforbruget pr. døråbning estimeret.

Energiforbruget til døråbninger kan bestemmes ved at benytte regressionslinien for energiforbrug. Ved denne metode elimineres behovet for at inddrage antal døråbninger og samlet døråbningstid. Denne metode vil give følgende beregninger: $((88,8 \text{ kWh/år} - 70 \text{ kWh/år}) / 88,8 \text{ kWh/år}) \times 100\% = 21\%$ svarende til et energiforbrug til døråbninger på 18,8 kWh/år eller 52 Wh/døgn.

Den andel af energiforbruget, der anvendes til døråbninger, kan estimeres til 19% for lavenergikøleskabet, 11% for referencekøleskabet (forudsat at energitabet ved døråbninger er uafhængig af typen på køleskabet) ud fra de forudsætninger, der er givet af Pedersen & Lawætz (1989). For et standardkøleskab årgang 1991 (årligt energiforbrug på 277 kWh jf. tabel 5.4.2), som indgår i undersøgelsen udført af Weidema *et al.* (1995), bliver andelen 5,4% for det gennemsnitlige køleskab i funktion i 1992 (årligt energiforbrug 345 kWh); andelen vil være faldende med alderen på køleskabene.

Dermed vil den største fejl, der kan begås ved at se bort fra antallet af døråbninger for det enkelte produkt være $\pm 5,4\%$.

Energiforbrug til nedkøling af fødevarer

Weidema et al. (1995) angiver følgende worst-case estimater:

- varmekapaciteten for fødevarer er $1,6 \text{ kJ}/^\circ\text{C}/\text{kg}$ (for energiforbruget til nedkøling er dette imidlertid ikke et worst-case estimat, men tværtimod et best-case estimat; et worst-case estimat ville være varmekapaciteten for vand $4,16 \text{ kJ}/^\circ\text{C}/\text{kg}$),
- andelen af madvarer, der skiftes/fyldes i køleskabet hver dag, er 25% ,
- temperaturforskæl mellem fødevarer og køleskabstemperatur er 15°C .

Følgende gennemsnits-estimater anvendes i stedet:

- varmekapaciteten for fødevarer: $2,9 \text{ kJ}/^\circ\text{C}/\text{kg}$ (CV: 28%),
- andelen af madvarer, der skiftes/fyldes i køleskabet hver dag: 14% (CV: 20%),
- temperaturforskæl mellem fødevarer og køleskabstemperatur: 10°C (CV: 5%).

Den gennemsnitlige fyldningsgrad for køleskabet er skønnet til 65%; se senere. Heraf kan det beregnes, at $200 \text{ liter} \times 0,65 \times 0,14 = 18,2 \text{ liter}$ udskiftes hver dag. Det forudsættes, at vægtfylden for fødevarerne er 1 kg/liter, hvorved energiforbruget til nedkøling af fødevarerne kan beregnes til: $18,2 \text{ kg/døgn} \times 10^\circ\text{C} \times 2,9 \text{ kJ}/^\circ\text{C}/\text{kg} = 528 \text{ kJ/døgn}$ svarende til $0,15 \text{ kWh/døgn}$ eller 32% af det samlede energiforbrug. Idet de nævnte tal antages at være uafhængige, kan variationskoefficienten på det samlede energiforbrug til nedkøling beregnes til 35%. Bidraget til usikkerheden på det samlede energiforbrug bliver da $32\% \times 35\% = 11\%$.

Udnyttelsesgrad for køleskab

Weidema et al. (1995) angiver et worst-case estimat for udnyttelsesgraden for køleskabe på 50%. Skønnet er baseret på en spørgeskemaundersøgelse af udnyttelsesgraden foretaget som et led i testen af lavenergi-køleskabet. Deltagerne i forsøget blev spurgt, om deres køleskab var 1) næsten fyldt, 2) halvt fyldt eller 3) kun lidt fyldt (Pedersen & Lawætz, 1989). 73 (64%) ud af 114 husstande svarede næsten fyldt, men forfatterne vurderede, at der ikke var grundlag for vurdering af sammenhængen mellem familiestørrelsen og fyldningsgraden. Kvaliteten af estimatet for udnyttelsesgraden kan angives som følger: 50% (4, 2, 1, 1, 1). Datakvalitetsindexet er vist i tabel 5.4.10.

Tabel 5.4.10

Datakvalitetsindex for skøn af udnyttelsesgrad for køleskabe.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed			
Pålidelighed	Der er tale om skøn og detaljeringsgraden/-præcisionen i spørgsmålene var lav.	4	CV: 15%, estimeret
Fuldstændighed	I spørgeskema undersøgelsen har der kun indgået 114 husstande.	2	CV: 3%, estimeret
Tidsmæssig korrelation	Mindre end 3 år.	1	
Geografisk korrelation	Dækende for Danmark.	1	
Teknologisk korrelation	Dækker de aktuelle køleskabe.	1	
Samlet usikkerhed			15%

Det giver ikke mening at angive usikkerheden på et worst-case estimat, idet der jo i dette allerede er taget højde for usikkerheden. Derimod kan den estimerede usikkerhed bruges til at angive en sandsynlig værdi, der altså her bliver $65\% \pm 15\%$.

Fortrængt energi til rumopvarmning

50% af varmetabet ved anvendelse af køleskabe er skønnet at fortrænge energi til opvarmning (Preben Buhl Pedersen, dk-Teknik, personlig kommunikation, 1993). Kvaliteten af data kan angives som $\pm 50\%$ (4, 5, 1, 1, 3-5); negativt fortægn udtrykker et negativt energiforbrug = energigevinst. Begrundelsen for det angivne datakvalitetsindex er vist i tabel 5.4.11.

Tabel 5.4.11

Datakvalitetsindex for skøn over fortrængt energi til rumopvarmning.

Datakvalitetsindikator	Datakvalitet	Score	Usikkerhed
Oprindelig usikkerhed			
Pålidelighed	Skønnet er foretaget af en ekspert, men det er ikke oplyst i hvor høj grad skønnet er baseret på faktisk viden.	4	CV: 10%, estimeret
Fuldstændighed	Kilden er angivet som personlig kommunikation, hvorfor det ikke er muligt at vurderre fuldstændigheden på skønnet.	5	CV: 25%
Tidsmæssig korrelation	Skønnet er aktuelt.	1	
Geografisk korrelation	Skønnet er dækende for danske forhold.	1	
Teknologisk korrelation	Fortrængt energi regnet som sparet el.	3-5	CV: 10%, estimeret
Samlet usikkerhed			21%

Vægtfylde for skinke

Vægtfylden for skinke er bestemt af producenten til $1,35 \text{ l/kg}$ (estimeret CV: 5%, (1, 1, 1, 1, 1)).

Opbevaringstid for skinke

Den gennemsnitlige opbevaringstid for skinke er skønnet af forfatterne til 7 dage (estimeret CV: 100%, (5, 1, 1, 1, 1)).

5.4.3 Samlet vurdering af datakvaliteten

Datakvaliteten for de indgående parametre er sammenfattet i tabel 5.4.12. Datakvaliteten er beskrevet i forhold til den funktionelle enhed, som er nedkøling af 1 kg skinke til den ønskede temperatur i én uge.

Datakvaliteten i det viste eksempel er meget svingende, idet nogle af de indgående data f.eks. omhandler forhold som udnyttelsesgrad for køleskabe og fortrængt energi til rumopvarmning. Disse forhold er ikke bestemt ved målinger, men derimod skønnede, hvorfor de optræder med et "dårligt" datakvalitets index.

Tabel 5.4.12

Middelværdi, variationskoefficient, datakvalitetsindex og total variationskoefficient for de indgående data.

Parameter	Værdi	CV	Datakvalitetsindex	Total CV
Salgstal	-	-	2.1.1.1.1	1%
Energiforbrug for køleskabe årgang 1991	277 kWh/år/køleskab	17%	2.1.1.1.1	18%
Konstant til bestemmelse af energiforbruget for ældre køleskabe	+ 10,5 kWh/år	17%	2.2.1.3.1	21%
Samlet energiforbrug for køleskabe i 1992	277 kWh/år + $T \times 10,5 \text{ kWh} = 769,410 \text{ kWh/år}$		Beregnet i afsnit 6.4	
Udnyttelsesgrad for køleskabe	50%		4.2.1.1.1	65% \pm 15%
Fortrængt energi til rumopvarmning	$\pm 50\%$		4.5.1.1.3-5	21%
Vægtfylde Skinke	1,35 l/kg	5%, estimeret	1.1.1.1.1	5%
Opbevaringstid for skinke	7 dage	100%, estimeret	5.1.1.1.1	100%

5.5 Sammenfattende vurdering af metoden til beskrivelse af datakvaliteten

En metode til beskrivelse af datakvaliteten i forbindelse med livscyklusvurdering er introduceret i afsnit 5.1. I afsnit 5.2 - 5.4 er metoden asprøvet på en række cases:

- Energiforbrug til landbrugsafgrøder
- Emissioner fra et bageri
- Energiforbrug til nedkøling af varer

Beskrivelsen af datakvaliteten indeholder i princippet tre trin:

- Bestemmelse af (den oprindelige) usikkerhed på de enkelte data (f.eks. variation på gennemsnit af flere målinger); usikkerheden kan bestemmes ved statistisk analyse af data eller estimeres.
- Beskrivelse af kvaliteten af de enkelte data ved scoring af data-kvalitetsindikatorernes pålidelighed, fuldstændighed, tidsmæssig, geografisk og teknologisk korrelation ud fra "datakvalitetsmatricen".
- Bestemmelse af de supplerende usikkerheder på de enkelte data i med dårlig score for en eller flere af datakvalitetsindikatorerne som estimerer eller bestemmer vha. andre datasæt; ud fra den oprindelige usikkerhed og de supplerende usikkerheder estimeres den samlede usikkerhed, der korrigerer for den dårlige datakvalitet.

Asprøvningen af den beskrevne metode på de tre cases giver anledning til en række kommentarer.

Korrigerede variationskoefficienter

De korrigerede variationskoefficienter, der medregner usikkerheden fra datakvaliteten, varierer i de forskellige præsenterede cases:

- fra < 1% (1,1,1,1,1) til 104% (5,5,5,5,3) for energiforbrug inden for landbrug
- fra < 1% (1,1,1,1,1) til 96% (1,1,1,2,1) for emissioner fra et bageri
- fra < 1% (2,1,1,1,1) til 21% (2,2,1,3,1) for energiforbrug til nedkøling af varer

Dominerende usikkerheder

Case 1 om energiforbrug inden for landbrug viser, at de korrigerede variationskoefficienter domineres af enkelte store usikkerheder. Ud fra denne observation synes det rimeligt at foreslå, at datakvalitetsindeksene primært benyttes som checkliste for identifikation af væsentlige usikkerheder, hvorefter der regnes videre med de væsentlige usikkerheder. Usikkerheden på det samlede resultat vil blive lidt undervurderet, hvilket kan imødegås ved introduktion af erfaringsbaserede korrektionsfaktorer. Uafhængigt af de supplerende usikkerheder kan der i visse tilfælde være tale om at justere middelværdierne som følge af viden om effektivisering af f.eks. energianvendelse til industrielle processer.

Beskrivelsens omfang

Anvendelse af den beskrevne metode i livscyklusvurderinger eller i andre sammenhænge, hvor der er behov for at beskrive datakvaliteten, kan synes overvældende. I praksis vil afsnittene "beskrivelse af casen" og "samlet vurdering af datakvaliteten" med den sammenfattende tabel kunne indgå i en rapport, mens det mellemliggende afsnit "beskrivelse af datakvaliteten" kunne placeres i bilag eller udelades. Alt efter de aktuelle datas kvalitet vil der være behov for en mere eller mindre detaljeret sammenfattende tabel, som det fremgår af case 2 om emissioner fra et bageri. Hvis datakvaliteten er dårlig, vil der være behov for at indicere, hvilke data der giver anledning til væsentlige supplerende usikkerheder med henblik på at forbedre datagrundlaget på netop disse punkter.

Detaljeringsniveau

Detaljeringsniveauet for beskrivelse af datakvalitet bør ligeledes fastsættes i overensstemmelse med formålet af den aktuelle undersøgelse, herunder hvem datakvaliteten skal dokumenteres overfor.

Entydigt datakvalitetsindex

Under arbejdet med beskrivelse af datakvaliteten har det vist sig, at de anvendte datakvalitetsindikatorer ikke i alle tilfælde entydigt beskriver data i de bearbejdede cases ligesom tildelingen af en score til de forskellige indikatorer ikke i alle tilfælde er entydig.

Den *første* situation kan illustreres med datakvalitetsindexet for emission af BOD og total-N, hvor der forekommer to spildevandsudløb med væsentlig forskellig koncentration af BOD og total-N (case 2). Forholdet mellem mængderne i de to udløb kendes ikke, og spørgsmålet er: hvilken datakvalitetsindikator beskriver det pågældende problem? - Er det et fuldkommenhedsproblem, eller er det et spørgsmål om geografisk korrelation? I den nævnte situation er det valgt at udtrykke datakvaliteten som et geografisk problem.

Den *anden* situation kan illustreres med scoringen af datakvaliteten for skønnet over fortrængt energi til rumopvarmning (case 3; tabel 5.4.11), hvor den teknologiske korrelation er tildelt scoren 3-5. Den ringe teknologiske korrelation kan begrundes i, at der er tale om at anvende spildvarme til opvarmning frem for varme produceret til formålet.

Supplerende usikkerheder

Tildeling af supplerende usikkerheder til de angivne datakvalitetsindex eller skøn over usikkerheder på det oprindelige data kan ligeledes være vanskelig. I visse tilfælde kan lignende undersøgelser anvendes til at estimere usikkerheden på oprindelige data med en passende grundlse, f.eks. vil variationen i energiforbruget til et givet køleskab kunne anvendes som skøn over variationen i energiforbruget for et gennemsnitskøleskab.

6. Statistisk analyse af usikkerheder

6.1 Introduktion til statistisk analyse

I nærværende kapitel beskrives statistisk analyse og resultat af 3 forskellige eksempler på dele af livscyklusvurderinger. Indledningsvist er de anvendte statistiske metoder beskrevet. I et enkelt eksempel er der endvidere anvendt simuleringer til analysen. Denne metode beskrives ligeledes indledningsvist.

6.1.1 Statistisk analyse

Livscyklusvurdering er i nærværende projekt foretaget ved at opstille et samlet flowforløb over hele processen. I denne proces indgår der en række enkeltpresser. Hver enkelt proces bidrager til den samlede livscyklusvurdering med en middelværdi, som der er en usikkerhed på. Den samlede livscyklusvurdering foretages på basis af de enkelte processers bidrag til et totalt niveau (middelværdi) og en total varians.

Som eksempel kan nævnes en simpel proces, hvor produktet C fås ved en kombination af A og B. Mængden af produktet C afhænger af mængden af såvel A som B. Usikkerheden på mængden af produktet C afhænger af usikkerheden på mængderne af A og B. Dette kan matematisk udtrykkes som følger:

$$\mu_C = \mu_A + \mu_B \quad (21)$$

$$\sigma_C = \sqrt{\sigma_A^2 + \sigma_B^2 - 2\sigma_A\sigma_B\rho} \quad (22)$$

hvor μ_X er middelværdi af X (hvor X kan være A, B eller C)

σ_X er standardafvigelsen på X

σ_X^2 er variansen på X

ρ er korrelationen (her mellem A og B)

Processer uafhængige

Det er i den statistiske analyse antaget, at de enkelte processer er uafhængige, dvs. korrelationen (ρ) antages at være nul, $\rho = 0$. Ovenstående udtryk for estimatet af standardafvigelsen på C forenkles herved til

$$\sigma_C = \sqrt{\sigma_A^2 + \sigma_B^2} \quad (23)$$

idet $\rho = 0$. Middelværdien af C er uændret.

Taylors approksimation

Det er ofte vanskeligt at finde eksakte udtryk for middelværdi og varians for mere komplicerede sammensatte funktioner (f.eks. produktet eller ratioen af to variable). Middelværdi og varians af sammensatte funktioner for 2 eller flere variable er bestemt ved

Taylors approksimation (se f.eks. Conradsen (1984)). Antages den sammensatte funktion at være givet ved:

$$h(X_1, X_2, \dots, X_n) \quad (24)$$

fås Taylor approksimationen omkring $\mu_i = E(X_i)$, $i = 1, \dots, n$

$$\begin{aligned} h(X_1, X_2, \dots, X_n) &\approx \\ h(\mu_1, \mu_2, \dots, \mu_n) + \sum_{i=1}^n &(X_i - \mu_i) \left(\frac{dh}{dx_i} \mid (X_1 = \mu_1, X_2 = \mu_2, \dots, X_n = \mu_n) \right) \end{aligned} \quad (25)$$

Middelværdi og varians for den sammensatte funktion fås som

$$E(h(X_1, X_2, \dots, X_n)) \approx h(\mu_1, \mu_2, \dots, \mu_n) \quad (26)$$

$$Var(h(X_1, X_2, \dots, X_n)) \approx \sum_{i=1}^n (Var(X_i) \left(\frac{dh}{dx_i} \mid (X_1 = \mu_1, \dots, X_n = \mu_n) \right))^2 \quad (27)$$

hvor der er forudsat uafhængighed mellem variablene X_1, X_2, \dots, X_n .

Ved anvendelse af Taylors approksimation er middelværdi og varians for den sammensatte funktion givet ved produktet af 2 variable (X₁, X₂)

$$\begin{aligned} h(X_1 X_2) &\approx h(\mu_1 \mu_2) + (X_1 - \mu_1) \left(\frac{dh}{dX_1} \mid (X_1 = \mu_1, X_2 = \mu_2) \right) \\ &+ (X_2 - \mu_2) \left(\frac{dh}{dX_2} \mid (X_1 = \mu_1, X_2 = \mu_2) \right) \\ &= \mu_1 \mu_2 + (X_1 - \mu_1) \mu_2 + (X_2 - \mu_2) \mu_1 \end{aligned} \quad (28)$$

$$E(h(X_1 X_2)) \approx h(\mu_1 \mu_2) = \mu_1 \mu_2 \quad (29)$$

$$\begin{aligned} Var(h(X_1 X_2)) &\approx Var(X_1) \left(\frac{dh}{dX_1} \mid (X_1 = \mu_1, X_2 = \mu_2) \right)^2 \\ &+ Var(X_2) \left(\frac{dh}{dX_2} \mid (X_1 = \mu_1, X_2 = \mu_2) \right)^2 \\ &= Var(X_1) \mu_2^2 + Var(X_2) \mu_1^2 \end{aligned} \quad (30)$$

Et eksempel på beregning af variansen for produktet af to eller flere variable er bestemmelsen af energiforbruget til produktion af pesticider anvendt i forbindelse med asgrøder til 1 kg brød:

$$E(Pest) = A \cdot B \cdot C \quad (31)$$

$$\begin{aligned} Var(Pest) &= Var(A \cdot B \cdot C) \\ &\approx Var(A)(\mu_B \mu_C)^2 + Var(B)(\mu_A \mu_C)^2 + Var(C)(\mu_A \mu_B)^2 \end{aligned} \quad (32)$$

- hvor
 A er anvendt mængde aktivt stof pr. ha dyrket areal (kg aktivt stof/ha dyrket areal)
 B er anvendt areal pr. kg brød (ha/kg brød)
 C er energiforbrug til produktion af 1 kg aktivt stof (MJ/kg aktivt stof)

Middelværdi og varians for ratioen mellem 2 variabler A og B er under antagelse af uafhængighed mellem variablerne givet ved

$$E\left(\frac{A}{B}\right) = \frac{\mu_A}{\mu_B} \quad (33)$$

$$Var\left(\frac{A}{B}\right) = \frac{1}{\mu_B^2} Var(A) + \frac{\mu_A^2}{\mu_B^4} Var(B) \quad (34)$$

Taylors approksimation er anvendt til beregning af middelværdi og varians for sammensatte funktioner, hvor eksakte udtryk kun vanskeligt kan bestemmes. Nøjagtigheden af Taylors approksimation afhænger af variationens størrelse på de eksakte variabler i den sammensatte funktion. Nøjagtigheden er størst, når variansen på de enkelte variabler er lille.

De approksimative formler for middelværdi og varians, givet ved Taylors approksimation, er ikke baseret på specifikke antagelser om typen af fordeling for de enkelte variabler.

Simuleringer

Til sammenligning er der foretaget beregning af middelværdi og varians ved simuleringer. Ved en simulering antages hver enkelt variabel i den sammensatte funktion at følge en specifik fordeling (f.eks. normalfordeling), med kendte (eller antagne) parametre i fordelingen (f.eks. middelværdi og varians i normalfordelingen).

Middelværdi og varians for sammensatte funktioner (f.eks. produktet XY) bestemmes i simulering ved at trække et tilfældigt tal fra hver af variablene respektive fordelinger. De udtrukne tal sammensættes til et tal i henhold til udtrykket for den sammensatte funktion. Udtagningen fra de respektive fordelinger og sammenregning til et tal gentages et antal gange (f.eks. 1000 gange). Middelværdi og varians for den sammensatte funktion bestemmes herefter som hhv. gennemsnit og varians af de gentagne udtrækninger og sammenregnede tal.

Som eksempel ses på produktet af 2 variabler XY. Det antages, at begge variabler følger en normalfordeling

$$X \in N(\mu_X, \sigma_X^2), \quad Y \in (\mu_Y, \sigma_Y^2) \quad (35)$$

Fra henholdsvis X og Y's fordelinger udtrækkes et tilfældigt tal som sammensættes ved produktet X_1Y_1 . Udtrækningen og multiplikationen fortsættes n gange (f.eks. n = 1000), hvorved haves: $X_1Y_1, X_2Y_2, \dots, X_nY_n$.

Middelværdi og varians bestemmes nu som

$$E(XY) = \frac{1}{n} \sum_{i=1}^n X_i Y_i \quad (36)$$

$$Var(XY) = \frac{1}{n-1} \sum_{i=1}^n (X_i Y_i - \bar{XY})^2 \quad (37)$$

Konfidensinterval

I livscyklusvurderingen angives middelværdi og varians for hver enkelt proces. Den samlede middelværdi (f.eks. energiforbruget) for hele processen bestemmes på baggrund af estimerer af middelværdierne for enkeltpresserne. Tilsvarende estimeres den totale varians på baggrund af enkeltpressernes varianser. Disse estimerer kan endvidere suppleres med konfidensintervaller, test af middelværdi, etc. Her angives blot et 95% konfidensinterval for en middelværdi:

$$\bar{X} \pm t(0,975;df) s \sqrt{\frac{1}{n}} \quad (38)$$

hvor \bar{X} er et estimat af middelværdien
 s er et estimat af standardafvigelsen
 n er antallet af observationer
 df er antallet af frihedsgrader (her, $df = n-1$)
 $t(0,975;df)$ er værdien fra en t-fordeling med df frihedsgrader svarende til 97,5% fraktilen

Middelværdi og standardafvigelse estimeres ved

$$\bar{x} = \frac{1}{n} \sum_{i=1}^n x_i \quad (39)$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^n (x_i - \bar{x})^2 \quad (40)$$

hvor x_i er observation i, $i=1, \dots, n$
 n er antal observationer

6.1.2 Anvendte edb-programmer

Der er foretaget indledende beregninger af forskellige deskriptive størrelser, såsom middelværdi, varians og standardafvigelse ved

anvendelse af traditionelle regnearksprogrammer f.eks. EXCEL (Microsoft, 1994) og Quattro Pro (Borland International, 1992). Den videre statistiske analyse og simulering er foretaget ved anvendelse af SAS (Statistical Analysis System) (SAS Institute, 1988). Simuleringer er endvidere foretaget ved anvendelse af @RISK (Palisade Corporation, 1991). EXCEL/Quattro Pro, SAS og @RISK er kort beskrevet herunder.

EXCEL/Quattro Pro

Regnearksprogrammerne EXCEL (Microsoft, 1994) og Quattro Pro (Borland International, 1992) fungerer alle stort set ud fra de samme principper, hvorfor beskrivelsen her ikke skelner mellem de forskellige programmer.

Regnearksprogrammerne er udviklet til behandling af større mængder af data, som er systematisk placeret i et "spreadsheet" (regnark) f.eks. som en serie målinger foretaget på forskellige tidspunkter. De udmærker sig ved, at de kan behandle data med alle almindelige regnearter, som kan være sammensat i en formel tilpasset det konkrete formål. Beregninger på måleserier (f.eks. koncentrationer af forskellige stoffer i spildevandsprøver udtaget f.eks. dagligt, ugentligt eller månedligt) udføres ved gentagne beregninger, hvilket i principippet gøres ved at kopiere/gentage de relevante formler.

Regnark indeholder de mest gængse statistiske funktioner som indlagte formler f.eks.:

- regressionsanalyse
- bestemmelse af middelværdien for et antal værdier
- bestemmelse af antal værdier
- bestemmelse af minimum og maximum værdien
- bestemmelse af summen af en række værdier
- bestemmelse af standardafvigelsen på middelværdien (af enten en stikprøve eller en population)
- bestemmelse af variansen på middelværdien (af enten en stikprøve eller en population)

Regnearksprogrammerne rummer ligeledes mulighed for at præsentere resultaterne i såvel tabeller som i grafer tilpasset det specifikke formål.

Regnearksprogrammerne er forsynet med "menu-linie", hvorunder funktioner kan aktiveres ved at klikke med musen på det ønskede menupunkt i et "rullegardin-menu" eller ved en tastekombination (f.eks. **ctrl +**). En række hyppigt anvendte funktioner kan være anbragt i en "værktøjslinie" evt. i form af ikoner. De fleste beregnings- og præsentationsfunktioner er således relativt let tilgængelige, uden at det er nødvendigt at kende teorien bag de forskellige funktioner.

SAS

Den statistiske analyse er foretaget ved anvendelse af den statistiske programpakke, SAS (Statistical Analysis System) (SAS Institute, 1988). SAS er nationalt og internationalt meget anvendt til en lang række statistiske analyser og problemstillinger.

Programpakken er opbygget med et antal moduler til forskellige kategorier af analyser. Som eksempel på moduler kan nævnes statistik, basis og grafikmoduler. Hvert modul er opbygget med en serie af procedurer, som kan anvendes til de statistiske analyser. Som eksempel kan nævnes procedurer til analyse af generelle lineære modeller, lineær og ikke-lineær regressionsanalyse, cluster-analyser, etc. Hver procedure (f.eks. til lineær regression) kan anvendes ved et kald af proceduren med specifikation af data og variabler, der skal benyttes, samt hvilke analyser m.m., der skal udføres. Der er endvidere mulighed for datahåndtering (omregning, transformationer, etc.) samt programmering af egne procedurer.

SAS fås til forskellige styresystemer, her kan bl.a. nævnes DOS, Windows og UNIX. Hele programpakken kræver megen diskplads. Afhængigt af styresystem og antal moduler, der er nødvendigt, kræves minimum ca. 50 - 100 Mb diskplads. For at få en effektiv afvikling af programmer kræves endvidere en del RAM, men er afhængigt af styresystem, mm.

SAS er en statistisk programpakke, der kræver en del indføring for at kunne benyttes effektivt. SAS er til gengæld en meget fleksibel programpakke, hvor i principippet alle typer problemer kan løses, idet der er mulighed for selv at programmere procedurer, som ikke allerede findes i programpakken. SAS kan i principippet benyttes til afvikling af uendelig store problemstillinger, begrænsningen ligger i computerens diskplads, RAM, mm.

I nærværende projekt er benyttet procedurer til beregning af deskriptive størrelser såsom middelværdi, standardafvigelse, varians (SAS Institute, 1990). Der er i vid udstrækning foretaget programmering til beregning af totale variationer, usikkerheder og variationskoefficienter for de samlede processer ud fra input data.

Tidsforbruget til de enkelte analyser er meget afhængigt af den enkeltes erfaring med SAS. For den rutinerede bruger tager analyserne kort tid, mens en ny bruger kan bruge mange timer på selv meget simple analyser.

@RISK

@RISK (Palisade Corporation, 1991) er et simuleringsværktøj, som kan anvendes i tilknytning til standardregnearksprogrammer som f.eks. EXCEL. Det betyder, at data indskrives i et regneark, og alle regnearkets muligheder for regneoperationer og grafisk fremstilling er til rådighed.

Grundprincippet i @RISK er, at alle data behandles som sandsynlighedsfordelinger (benævnt "inputvariable"), som karakteriseres ved f.eks. en middelværdi og en standardafvigelse. Der er 28 forskellige sandsynlighedsfordelinger til rådighed i programmet. I den samme beregning kan indgå inputvariabler, som er repræsenteret med forskellige fordelinger, dvs. man er ikke begrænset af, at alle inputvariabler repræsenteres med f.eks. normalfordelinger. Fordelingsfunktionerne anvendes i regnearket på samme måde som andre funktioner, og sandsynlighedsfordelingernes parametre kan repræsenteres med cellehenvisninger.

Som eksempel vil energiforbruget til produktion af pesticider (jf. tabel 5.1.5) kunne repræsenteres som en normalsfordeling med en middelværdi på 288 MJ/kg og en standardafvigelse på 179 MJ/kg (CV: 62%). På lignende måde kan pesticidforbruget pr. ha for danske afgrøder (tabel 5.2.6) repræsenteres. Energiforbruget pr. ha til fremstilling af pesticider kan beregnes ved at multiplicere de 2 tilfældige tal udtrukket fra de 2 respektive fordelinger. Umiddelbart vil beregningen i regnearket kun ses som produktet af funktionernes middelværdier. Det er muligt at foretage en simulering, hvor beregningen - i dette eksempel multiplikationen af de to inputvariabler - gentages mange gange. Ved hver beregning udtages der fra hver af de to fordelinger ved hjælp af en tilfældighedsgenerator en tilfældig værdi, hvorefter disse værdier multipliceres. De mange beregnede værdier præsenteres i @RISK som en fordeling, der beskrives med middelværdi, varians, fraktiler osv. Der er ligeledes mulighed for grafisk afbildning af fordelingen.

I beregningerne kan indgå en lang række inputvariabler, og antallet af operatorer er i principippet ubegrænset. I denne form for analyse vil korrelationer mellem de enkelte inputvariabler kunne have en væsentlig indflydelse på beregningen, og det er muligt at opstille en korrelationsmatrix, der angiver, i hvilken grad der er positiv eller negativ korrelation mellem de enkelte input variabler.

Efter simuleringen kan der foretages forskellige statistiske analyser på resultaterne. Ved beregning af korrelationer mellem de forskellige inputvariabler og outputvariablen kan der på en enkelt måde laves en følsomhedsanalyse, der giver en rangordning af hvilke inputvariabler, der har størst indflydelse på standardafvigelsen af outputvariablen. Analysen viser hvilke data, det er værd at forsøge at bestemme mere præcist og hvilke data, der kun har marginal indflydelse på usikkerheden på det beregnede resultat.

Programmet er relativt enkelt at bruge, og det har den fordel, at der i principippet ikke er begrænsninger for hvilke regneoperationer, der opstilles i regnearket, og at man har alle EXCELS værktøjer til rådighed. Svaghederne er, at man selv skal opbygge skemaer og beregningsfunktioner i regnearket, og ikke umiddelbart kan se indtastede data og relationer i form af "flow-charts". Der er altså tale om et værktøj til den professionelle bruger, der arbejder med opgaver, som ikke passer til færdigudviklet LCA-software.

Programmet kræver for at kunne køre EXCEL mindst 8 Mb RAM og ideelt set 16 Mb, hvis man ikke hurtigt skal støde ind i begrænsninger.

Man skal være opmærksom på, at simulering ikke er en eksakt metode. Det betyder, at f.eks. middelværdien vil være forskellig fra simulering til simulering. Med et stigende antal gennemregninger vil disse forskelle imidlertid mindskes. Det er muligt at få programmet til automatisk at fortsætte gennemregningerne indtil ændringerne i forskellige outputparametre er mindre end en givet procentdel. Simuleringen skal ikke forveksles med den praksis, at man tager et stigende antal uafhængige prøver for at bestemme en størrelse så præcist som muligt.

Øget antal gennemregninger vil ikke mindske standardafvigelsen på middelværdien, men blot bestemme standardafvigelsen mere præcist.

Der er muligt i @RISK at standse simuleringen, når ændringer i outputvariablen målt over et vist antal gennemregninger er under en vis grænse i stedet for at definere et antal gennemregninger. Metoden er dog ikke særlig veldefineret i forhold til statistiske termer. I mere avancerede programmer til arbejde med sikkerhedsanalyse som f.eks. Comrel (RCP-GmbH, München) er det muligt at beskrive sikkerheden på en undersøgt outputparameter f.eks. standardafvigelsen på outputvariablen med en variationskoefficient på estimatet. Vi vil ikke komme nærmere ind på de statistiske forhold vedrørende dette, men blot konstatere at dette ikke er muligt i @RISK, hvor usikkerheden på estimatet kun kan bestemmes ved at foretage flere simuleringer og sammenligne resultaterne.

I det beregnede eksempel "Energiforbrug til landbrugsafgrøder" indgår der mere end 100 inputvariabler. Med 1000 gennemregninger giver beregningen reproducerbar standardafvigelse på outputvariablen. For at kunne foretage en fejlanalyse er det nødvendigt, at alle fordelinger af inputvariabler gemmes i computerens hukommelse. Her vil der meget nemt kunne opstå begrænsninger, der gør det mest hensigtsmæssigt at køre en selvstændig simulering med færre gennemregninger for denne del af beregningen. Rangordningen af inputvariabler var i det gennemregnede eksempel selv med 300 gennemregninger helt reproducerbar. Ved større regneoperationer er det muligt at opdele beregningen, således at outputvariabler fra én beregning anvendes som inputvariabler i en anden beregning.

6.2 Case 1: Energiforbrug til landbrugsafgrøder

I nærværende afsnit beskrives den statistiske analyse foretaget af Case 1: energiforbrug til landbrugsafgrøder. Selve casen er beskrevet i afsnit 5.2.1. Data og datakvalitet er beskrevet i afsnit 5.2.2. Disse 2 afsnit er udgangspunktet for den nedenfor beskrevne statistiske analyse.

Til beregning af energiforbruget til produktion af et brød på 1 kg er følgende processer medregnet mht. anvendelse af energi:

- produktion af pesticider
- produktion af N, P, K
- produktion af kalk
- forbrug af husdyrgødning
- markarbejde
- tørring af afgrøder

Samtlige beregninger er foretaget for 1 kg brød.

Der er foretaget beregning af middelværdi og varians for samtlige processer ved anvendelse af såvel Taylors approksimation som simulation i programmet SAS. I simulationen er der anvendt normalfordelinger for samtlige inputvariabler. For at undgå problemer med

udtrækning af negative tal i simulationerne er der endvidere anvendt en antagelse om logaritmisk normalsfordeling. I simulationerne er anvendt samme middelværdi og varians som ved Taylors approksimation for alle inputvariabler. Simuleringen er endvidere i 6.2.2 gennemført med programmet @RISK.

6.2.1 Statistisk analyse af data

Anvendt areal til produktion af afgrøderne til 1 kg brød (ha/kg brød)

Rugbrødet består af 8 afgrøder med sammensætning som angivet i tabel 5.2.2; høstudbyttet for de 8 afgrøder fremgår af tabel 5.2.3. Det anvendte areal til produktion af 1 kg rugbrød beregnes på baggrund af rugbrødets mængdesammensætning og høstudbyttet af de enkelte afgrøder, som følger

$$E(\text{areal}) = \frac{X}{Y} \quad (41)$$

hvor X er mængde afgrøde til 1 kg brød (kg afgrøde/kg brød)

Y er udbytte af afgrøde (kg afgrøde/ha)

Herved bliver anvendt areal samt usikkerhed for de enkelte afgrøder til 1 kg brød, som angivet i tabel 6.2.1.

Tabel 6.2.1

Anvendt areal til 1 kg rugbrød samt standardafvigelse og variationskoefficient.

Afgrøde	Mængde afgrøde til 1 kg brød ¹ (kg/kg)	Udbytte af afgrøde ² (kg/ha)	Anvendt areal til 1 kg brød (ha/kg)	Standardafvigelse (ha/kg)	Variationskoefficient %
Rug	0,536	4.460	0,00012018	0,000030069	24,9%
Sukkerroer	0,286	49.600	0,00000577	0,000000901	15,5%
Hvede	0,122	6.917	0,00001764	0,000002300	13,1%
Hørfrø, UK	0,015	1.683	0,00000891	0,000001607	18,0%
Hørfrø, Rus.	0,015	209	0,00007177	0,000014372	20,1%
Palmeolie	0,032	22.500	0,00000142	0,000000383	28,6%
Sesamfrø	0,004	556	0,00000719	0,000001584	22,2%
Solsikkefrø	0,030	1.345	0,00002230	0,000004467	20,2%
Totalt	1,040		0,00025519	0,000033793	13,2%

1. Fra tabel 5.2.2.

2. Fra tabel 5.2.3.

Energiforbrug til produktion af pesticider til 1 kg brød

Energiforbrug til produktion af pesticider til afgrøder til 1 kg brød fås som produktet af anvendt mængde pesticider (kg aktivt stof/ha dyrket areal) (pesticidsforbruget er angivet til 2 kg/ha for danske såvel som udenlandske afgrøder med en usikkerhed på henholdsvis 37% og 71% (CV)), anvendt areal pr. brød (ha/kg brød) og energiforbrug til produktion af 1 kg aktivt stof (MJ/kg aktivt stof) (energiforbruget til produktion af pesticider er beregnet til 288 MJ/kg (CV: 62%), ved

$$E(Pest) = A \cdot B \cdot C \quad (42)$$

- hvor A er anvendt mængde aktivt stof pr. ha dyrket areal (kg aktivt stof/ha dyrket areal); se tabel 5.2.5
 B er anvendt areal pr. kg brød (ha/kg brød); se tabel 6.2.1
 C er energiforbrug til produktion af 1 kg aktivt stof (MJ/kg aktivt stof)

Energiforbrug til produktion af kvælstof (N) til 1 kg brød
 Energiforbrug til produktion af kvælstof (N) til afgrøder til 1 kg brød fås som

$$E(N) = A \cdot B \cdot C_1 \cdot D \cdot E + A \cdot B \cdot C_2 \cdot D \cdot (1 - E) \quad (43)$$

- hvor A er anbefalet mængde kvælstof til de enkelte afgrøder (kg gødning/ha dyrket areal); se tabel 5.2.10
 B er anvendt areal pr. kg brød (ha/kg brød); se tabel 6.2.1
 C₁ er energiforbrug til produktion af 1 kg kvælstof fra urea (MJ/kg kvælstof); se tabel 5.2.14
 C₂ er energiforbrug til produktion af 1 kg kvælstof fra ammonium-nitrat mm. (MJ/kg kvælstof); se tabel 5.2.13/5.2.15
 D er dækningsgraden; se tabel 5.2.10
 E er andelen af kvælstof produceret fra urea; se tabel 5.2.16

Energiforbruget til produktion af N-kunstgødning anvendt i Danmark er estimeret til 38 MJ/kg N (CV: 10%); se tabel 5.2.15. Energiforbruget til produktion af N-kunstgødning anvendt i udlandet er estimeret til 46 MJ/kg N (CV: 16%) for ammoniumnitrat og komplekse gødninger (se tabel 5.2.13) henholdsvis 63 MJ/kg N (CV: 11%) for fremstilling af urea (se tabel 5.2.14).

Energiforbrug til produktion af phosphor (P) til 1 kg brød
 Energiforbrug til produktion af phosphor (P) til afgrøder til 1 kg brød fås som

$$E(P) = A \cdot B \cdot C \cdot D \quad (44)$$

- hvor A er anbefalet mængde phosphor til de enkelte afgrøder (kg gødning/ha dyrket areal); se tabel 5.2.10
 B er anvendt areal pr. kg brød (ha/kg brød); se tabel 6.2.1
 C er energiforbrug til produktion af 1 kg phosphor (MJ/kg phosphor); se tabel 5.2.17
 D er dækningsgraden; se tabel 5.2.10

Energiforbruget til fremstilling af phosphorgødning er estimeret til 28,3 MJ/kg P (CV: 15%); jf. tabel 5.2.17.

Energiforbrug til produktion af kalium (K) til 1 kg brød
 Energiforbrug til produktion af kalium (K) til afgrøder til 1 kg brød fås som

$$E(K) = A \cdot B \cdot C \cdot D \quad (45)$$

- hvor A er anbefalet mængde kalium til de enkelte afgrøder (kg gødning/ha dyrket areal); se tabel 5.2.10
 B er anvendt areal pr. kg brød (ha/kg brød); se tabel 6.2.1
 C er energiforbrug til produktion af 1 kg kalium (MJ/kg kalium)
 D er dækningsgrad; se tabel 5.2.10

Energiforbruget til fremstilling af K-gødning er opgivet til 5 MJ/kg K (CV: 29%); jf. tabel 5.2.18.

Energiforbrug til produktion af kalk ($CaCO_3$) til 1 kg brød
 Energiforbrug til produktion af kalk ($CaCO_3$) til afgrøder til 1 kg brød fås som

$$E(Kalk) = A \cdot B \cdot C \quad (46)$$

- hvor A er anbefalet mængde kalk til de enkelte afgrøder (kg gødning/ha dyrket areal)
 B er anvendt areal pr. kg brød (ha/kg brød); se tabel 6.2.1
 C er energiforbrug til produktion af 1 kg kalk (MJ/kg kalk)

Energiforbruget til fremstilling af kalk er estimeret til 0,8 MJ/kg (CV: 104%) jf. tabel 5.2.20, og der er regnet med et forbrug af kalk på 790 kg/ha (CV: 10,3%) til kalkkrævende danske afgrøder (sukkerroer og hvede), jf. tabel 5.2.19.

Energiforbrug til produktion af husdyrgødning til 1 kg brød
 Energiforbrug til produktion af husdyrgødning til afgrøder til 1 kg brød fås som

$$E(N-husdyr) = A \cdot B \cdot C \cdot \frac{(1 - D)}{E} \quad (47)$$

- hvor A er anbefalet mængde kvælstof til de enkelte afgrøder (kg gødning/ha dyrket areal); se tabel 5.2.10
 B er anvendt areal pr. kg brød (ha/kg brød); se tabel 6.2.1
 C er energiforbrug til produktion af 1 kg kvælstof i husdyrgødning (MJ/kg kvælstof)
 D er dækningsgraden; se tabel 5.2.10
 E er udnyttelsesgraden i husdyrgødning

Energiforbrug til markarbejde til 1 kg brød
 Energiforbrug til markarbejde til produktion af afgrøder til 1 kg brød fås som

$$E(Mark) = B \cdot C \quad (48)$$

- hvor B er anvendt areal pr. kg brød (ha/kg brød); se tabel 6.2.1
 C er energiforbrug pr. ha dyrket areal (MJ/ha); se tabel 5.2.23, hvor energiforbruget til markarbejde for udenlandske afgrøder er præsenteret

Energiforbruget til markarbejde i Danmark er 2.078 MJ/ha, 3.767 MJ/ha og 2.552 MJ/ha for henholdsvis rug, sukkerroer og hvede.

Energiforbrug til tørring af afgrøder til 1 kg brød

Energiforbrug til tørring af afgrøder til produktion af 1 kg brød fås som

$$E(Tør) = X \cdot C \quad (49)$$

hvor X er mængde afgrøde, der skal tørres til 1 kg brød (kg afgrøde/kg brød)

C er energiforbrug til tørring af 1 kg afgrøde (MJ/kg afgrøde)

Energiforbruget til tørring af korn er estimeret til 0,22 MJ/kg (CV: 25%), jf. tabel 5.2.24.

Det samlede energiforbrug til produktion af et brød på 1 kg

Tabel 6.2.2 viser det samlede energiforbrug beregnet ved anvendelse af Taylors approksimationer. Tabellerne 6.2.3 og 6.2.4 viser tilsvarende det samlede energiforbrug bestemt ved simulationer. Der er i begge tilfælde anvendt 5000 simulationer. Inputvariablerne er i tabel 6.2.3 antaget at følge normalfordelinger og i tabel 6.2.4 antaget at følge logaritmiske normalfordelinger.

Tabel 6.2.2

Energiforbrug til produktion af råvarer til rugbrød samt standardafvigelse og varians bestemt ved Taylors approksimationer.

Afgrøde	Energi (middelværdi) MJ/kg	Standard- afvigelse MJ/kg	Varians (MJ/kg) ²	%andel af total varians	Variations- koefficient %
Rug	1,05987	0,25985	0,0675	76	24,5
Sukkerroer	0,05728	0,01159	0,0001	0,2	20,1
Hvede	0,21027	0,03311	0,0011	1,2	15,7
Hørfrø, UK	0,11920	0,05401	0,0029	3,3	45,3
Hørfrø, Rusl.	0,32667	0,12221	0,0149	17	37,4
Palmeolie	0,01451	0,00448	0,0000	0,0	30,3
Sesamfrø	0,04057	0,01732	0,0003	0,3	42,7
Solsikkefrø	0,11322	0,04388	0,0019	2,2	38,7
Totalt	1,94159	0,29808	0,0888	100	15,3

Det samlede energiforbrug til produktion af råvarer til fremstilling af et brød på 1 kg er bestemt til 1,9 MJ med en standardafvigelse på 0,3 MJ ved Taylor approksimationer. Energiforbruget kan ligeledes udtrykkes ved et 95% konfidensinterval som $1,9 \text{ MJ} \pm 0,6 \text{ MJ}$ eller som $1,3 - 2,5 \text{ MJ}$. Ved simulering med normalfordelinger er energiforbruget bestemt til 2,1 MJ med en standardafvigelse på 0,41 MJ. Den tilsvarende simulering med logaritmiske normalfordelinger giver et samlet energiforbrug på 2,0 MJ og en standardafvigelse på 0,32 MJ.

Tabel 6.2.3

Energiforbrug til produktion af råvarer til rugbrød samt standardafvigelse og varians bestemt ved simulation¹ i normalfordeling (n = 5000).

Afgrøde	Energi (middelværdi) MJ/kg	Standard- afvigelse MJ/kg	Varians (MJ/kg) ²	%andel af total varians	Variations- koefficient %
Rug	1,13420	0,37457	0,1403	83	33,0
Sukkerroer	0,05816	0,01189	0,0001	0,1	20,3
Hvede	0,21427	0,03449	0,0012	0,7	16,1
Hørfrø, UK	0,12500	0,05908	0,0035	2,1	47,2
Hørfrø, Rusl.	0,34357	0,14408	0,0207	12	41,9
Palmeolie	0,01571	0,00619	0,0000	0,0	38,9
Sesamfrø	0,04323	0,01989	0,0004	0,2	45,8
Solsikkefrø	0,11680	0,04904	0,0024	1,4	42,0
Totalt	2,05092	0,41076	0,1687	100	20,0

1. Simuleringen er foretaget ved en Box-Muller transformation af uniformt fordelte tilfældige tal.

Tabel 6.2.4

Energiforbrug til produktion af råvarer til rugbrød samt standardafvigelse og varians bestemt ved simulation¹ i log-normalfordeling (n = 5000).

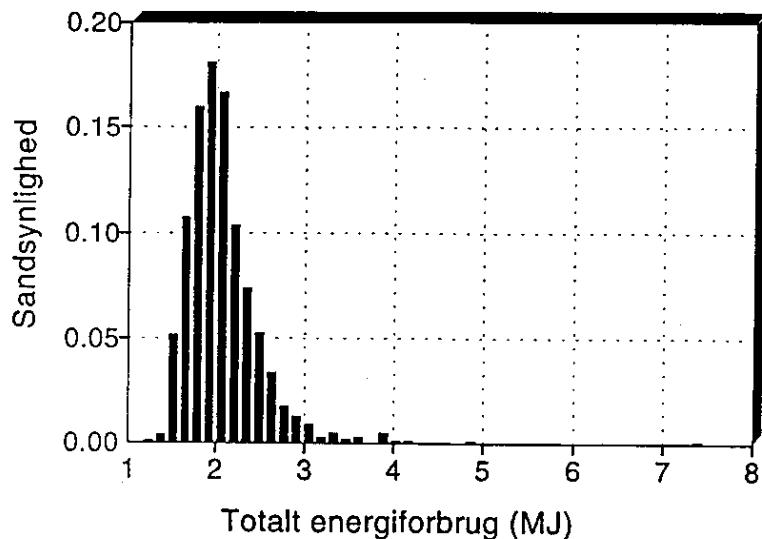
Afgrøde	Energi (middelværdi) MJ/kg	Standard- afvigelse MJ/kg	Varians (MJ/kg) ²	%andel af total varians	Variations- koefficient %
Rug	1,11887	0,27601	0,0762	76	24,7
Sukkerroer	0,05802	0,01147	0,0001	0,1	19,7
Hvede	0,21386	0,03340	0,0011	1,1	15,6
Hørfrø, UK	0,12437	0,05863	0,0034	3,4	47,1
Hørfrø, Rusl.	0,33884	0,12900	0,0166	17	38,1
Palmeolie	0,01540	0,00468	0,0000	0,0	29,9
Sesamfrø	0,04266	0,01836	0,0003	0,3	43,0
Solsikkefrø	0,11658	0,04753	0,0023	2,3	40,8
Totalt	2,02858	0,31643	0,1001	100	15,6

1. Simuleringen er foretaget ved en Box-Muller transformation af uniformt fordelte tilfældige tal.

6.2.2 Simulering af data med @RISK

Energiforbruget til produktion af rugbrød på 1 kg er i nærværende afsnit beregnet ved simulering i @RISK. I figur 6.2.1 er resultatet af simulering med @RISK illustreret med fordelingen af 5.000 gennemregninger af det totale energiforbrug til produktion af råvarer til fremstilling af 1 kg rugbrød. Alle inputvariabler er antaget at følge normalfordelinger uden korrelation mellem inputvariablerne.

Med 5.000 gennemregninger varierede standardafvigelsen på outputvariablen mindre end 1% ved gentagene simuleringer.



Figur 6.2.1

Fordeling af 5.000 gennemregninger af det samlede energiforbrug til produktion af råvarer til fremstilling af 1 kg rugbrød simuleret med @Risk. Simuleringen er foretaget med "Latin Hypercube sampling".

Som det fremgår af figur 6.2.1, er fordelingen af outputvariablen "totalt energiforbrug" ikke symmetrisk. Dette skyldes, at der i beregningen indgår bl.a. multiplikation og division med normalsfordelinger, hvorfor man heller ikke umiddelbart kan sige noget om, hvilken fordeling det er. Den asymmetriske fordeling af outputvariablen bevirker, at middelværdien i fordelingen er større end den værdi, som fremkommer, hvis det totale energiforbrug beregnes på baggrund af middelværdierne af inputvariablerne. Middelværdien i den angivne fordeling er 2,049, mens værdien beregnet på grundlag af middelværdier er 1,941 (se tabel 6.2.6 og 6.2.7). Den skæve fordeling betyder også, at den mest sandsynlige værdi (højeste punkt på fordelingen) ikke vil være middelværdien, men være lidt lavere. Det kan intuitivt synes forkert, at middelværdien ikke vil være den samme, men det er et resultat af det generelle fænomen, at den inverse fordeling af en symmetrisk fordeling normalt ikke er symmetrisk.

Simuleringen beregner ikke blot usikkerheden på resultatet, men også middelværdien på denne størrelse. Konsekvensen af dette har stor betydning, fordi det indebærer, at man ikke umiddelbart kan betragte en værdi beregnet på grundlag af middelværdier (jf. beregninger udført i forbindelse med livscyklusvurderinger generelt) som den korrekte middelværdi af resultatet.

Simuleringsprogrammet giver en række statistiske oplysninger om outputvariablen, hvoraf et udsnit fremgår af tabel 6.2.5. Som det ses, vil

90% af resultaterne ligge inden for intervallet 1,56 - 2,74 MJ/kg (intervallet fra 5% til 95% fraktilen). Da fordelingen ikke er symmetrisk, kan resultatet ikke præsenteres fyldestgørende med kun middelværdi og standardafvigelse, men bedre med middelværdi og relevante fraktiler.

Tabel 6.2.5

Statistiske oplysninger vedrørende det samlede energiforbrug (MJ/kg) til produktion af råvarer til fremstilling af rugbrød.

Middelværdi	2,05
Minimum	1,22
Maximum	10,77
Standardafvigelse, σ	0,43
Varians, σ^2	0,18
Skævhed	4,05
Kurtosis	56,41
5% fraktil	1,56
95% fraktil	2,74

I tabel 6.2.6 og 6.2.7 er total energiforbrug til produktion af 1 kg rugbrød opdelt på afgrøder henholdsvis processer. Standardafvigelsen på det totale energiforbrug er dels angivet som standardafvigelsen bestemt ved simulering, dels beregnet med Gauss fejlphobningslov (jf. afs. 4.2.1) ud fra standardafvigelsen på bidragene fra de enkelte afgrøder henholdsvis processer. Som det fremgår af tabel 6.2.7, er standardafvigelsen på det totale energiforbrug bestemt ved simulering på enkeltafgrøder større end standardafvigelsen beregnet på processer. Årsagen er, at der ved beregningen af energiforbrug til flere af processerne indgår samme inputvariabler, og at der således vil være korrelation mellem mellemresultaterne, såkaldt "funktionelle afhængigheder". Det samme er ikke tilfældet, hvis samme beregning foretages på baggrund af standardafvigelsen på de samlede bidrag fra afgrøderne (tabel 6.2.6). I dette tilfælde er standardafvigelsen den samme, fordi der til beregningen af de enkelte afgrøders bidrag ikke anvendes samme inputvariabler.

Tabel 6.2.6

Total energiforbrug til produktion af råvarer til fremstilling af rugbrød samt standardafvigelse og variationskoefficient relateret til afgøder bestemt ved simulation med @RISK.

Afgøde	Energiforbrug middelværdi (MJ/kg) μ	Energiforbrug simuleret	Standard- afvigelse	Variations- koefficient %
Rug	1,060	1.136	0,391	34,4
Sukkerroer	0,057	0,058	0,013	22,8
Hvede	0,210	0,214	0,037	17,5
Hørfrø, UK	0,119	0,123	0,059	47,6
Hørfrø, Rusland	0,327	0,342	0,139	40,7
Palmeolie	0,015	0,015	0,003	23,0
Sesamfrø	0,041	0,043	0,021	49,4
Solsikkefrø	0,113	0,118	0,049	41,0
I alt	1,941	2,049	0,425	20,7
Standardafvigelse beregnet ud fra mellemresultater			0,424	

Tabel 6.2.7

Total energiforbrug til produktion af råvarer til fremstilling af rugbrød samt standardafvigelse og variationskoefficient relateret til processer.

Proces	Energiforbrug middelværdi (MJ/kg) μ	Energiforbrug simuleret	Standard- afvigelse	Variations- koefficient %
Pesticider	0,147	0,156	0,083	53,0
N	1,036	1,100	0,273	24,8
P	0,095	0,100	0,028	28,4
K	0,033	0,035	0,010	27,5
Kalk	0,015	0,015	0,013	85,9
Husdyrgødning	0,003	0,003	0,002	66,0
Markarbejde	0,470	0,499	0,149	29,9
Tørring	0,142	0,142	0,030	21,0
I alt	1,941	2,049	0,425	20,7
Standardafvigelse beregnet ud fra mellemresultater			0,325	

Normalt vil man ved simulering aldrig beregne standardafvigelsen ud fra delresultater, men i komplekse tilfælde med mange parametre kan det muligvis være hensigtsmæssigt. Som det fremgår, skal man i disse tilfælde være opmærksom på korrelationer mellem mellemresultater.

6.2.3 Vurdering af resultaterne af den statistiske analyse af data

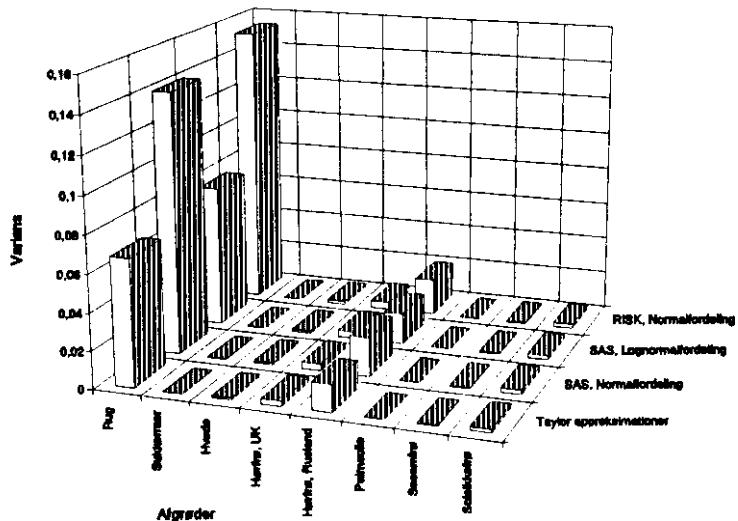
Den statistiske analyse af casen viser, at det samlede energiforbrug er estimeret til 1,94 MJ/kg brød med en standardafvigelse på 0,30 MJ/kg brød ved anvendelse af Taylors approksimationer. Anvendes

derimod simulation er det samme energiforbrug bestemt til 2,05 MJ/kg brød med en usikkerhed 0,41 MJ/kg brød ved antagelse af en normalfordeling for alle inputvariabler. Antages derimod, at alle inputvariabler er logaritmisk normalsfordelte, er det samlede energiforbrug bestemt til 2,03 MJ/kg brød med en usikkerhed på 0,32 MJ/kg brød. Anvendes @RISK til simulationen med antagelse om normalfordeling for alle inputvariabler fås et samlet energiforbrug på 2,05 MJ/kg pr. brød med en standardafvigelse på 0,43 MJ/kg pr. brød. Resultaterne er sammenfattet i tabel 6.2.8 og figur 6.2.2, hvor variansen for hver proces er angivet i procent af den samlede varians bestemt ved Taylors approksimationer og simulationer med @RISK og SAS.

Tabel 6.2.8

Samlet oversigt over det samlede energiforbrug og standardafvigelsen bestemt ved de 4 forskellige metoder.

Metode	Fordelingsantagelser	Energiforbrug MJ/kg	Standardafvigelse MJ/kg
Taylors approksimation	ingen	1.94	0.30
Simulation (SAS)	normalfordeling	2,05	0.41
Simulation (SAS)	lognormalfordeling	2.03	0,32
Simulation (@RISK)	normalfordeling	2.05	0.43



Figur 6.2.2

Variansen for energiforbruget til produktion af 1 kg rugbrød bestemt ved Taylors approksimationer og simulationer med @RISK og SAS.

Taylors approksimationer er benyttet til beregning af middelværdi og standardafvigelse for det totale energiforbrug. Approksimationerne kan generelt anvendes uafhængigt af fordelingstype, og den beregnede middelværdi og standardafvigelse beror dersør ikke på bestemte forde-

lingsantagelser. Der skal dog gøres opmærksom på, at der er tale om approksimationer, der har størst nøjagtighed for stokastiske variabler med et lille variationsområde (lille varians).

Simulationerne er baseret på bestemte antagelser om inputvariabernes fordelinger. Der er anvendt antagelser om såvel normalfordelinger som logaritmisk normalfordeling. Endvidere er simulationerne foretaget ved såvel programmering i SAS som ved anvendelse af @RISK.

Antagelserne om logaritmisk normalfordeling resulterer i et samlet energiforbrug og en standardafvigelse, der er tæt på de tilsvarende størrelser bestemt ved Taylors approksimationer. Ved anvendelse af logaritmiske normalfordelinger fremfor normalfordelinger opnås endvidere, at der ikke udtrækkes negative tal, der kan resultere i negative energiforbrug.

Antagelsen om normalfordelinger resulterer i større standardafvigelser på energiforbruget end approksimationerne og antagelsen om logaritmisk normalfordeling af inputvariablerne. Resultaterne ved anvendelse af @RISK og SAS ses at være i god overensstemmelse.

Konklusion

Anvendelse af SAS/Taylor approksimationer og @RISK og SAS-simuleringer til bestemmelse af den samlede usikkerhed på outputdata fra en udvalgt del af en livscyklusvurdering peger på, at SAS/Taylor approksimationer er håndterbart for mindre datasæt, mens simuleringer (@RISK og SAS) er bedre til at håndtere større datasæt.

6.2.4 Anvendelse af statistisk analyse livscyklusvurdering

Ved en videre analyse af data illustreres anvendelsen af statistisk analyse til at belyse en række generelle problemstillinger inden for livscyklusvurdering:

- Ved en *fejlanalyse* illustreres, hvilke usikkerheder på inputvariablerne der er af størst betydning for den samlede usikkerhed.
- Usikkerheden på det samlede resultat anvendes til at vurdere mulighederne for at udelade visse led i de samlede beregninger eller til at pege på led, som det ikke kan "betale sig" at bestemme mere nøjagtigt, da det ikke vil have væsentlig indflydelse på den samlede usikkerhed (*cut off*¹).
- Ved sammenligning af usikkerheden på energiforbruget for det aktuelle brød med usikkerheden på energiforbruget til fremstilling af et hypotetisk brød illustreres princippet om inddragelse af usikkerheder ved *sammenligning af to forskellige produkter* eller et eksisterende produkt med et forbedret produkt.

¹ Der eksisterer ikke nogen decideret definition af "cut off", men formålet kan være at begrænse sin dataindsamling ved at undlade at indsamle data for fx råvareproduktion, hvis de pågældende råvarer indgår med begrænset andel af det færdige produkt. For en mere detaljeret diskussion af denne problematik henvises til fx Lindfors *et al.* (1995).

På baggrund af simuleringen kan der i @RISK laves en fejlanalyse, hvor der analyseres for sammenhænge mellem outputvariablen (det totale energiforbrug) og de enkelte inputvariabler. I programmet analyseres der automatisk for sammenhænge mellem inputvariabler og outputvariablen ved multipel regression og "Spearman rank korrelation".

I tabel 6.2.9 er følsomhederne mellem de enkelte inputvariabler og outputvariablen rangordnet for de 10 inputvariabler med højeste følsomhed (positiv eller negativ). Følsomheden er angivet ved en følsomhedskoefficient. Denne koefficient angiver antallet af standardafvigeler, som outputvariablen øges ved at ændre standardafvigelsen på de enkelte inputvariabler med en standardafvigelse (for yderligere detaljer, se f.eks. Winston (1996)). Følsomhedskoefficienten og Spearmans Rank korrelationskoefficient angiver, hvor meget af variationen på outputvariablen, der kan forklares med et lineært forhold mellem inputvariablerne og outputvariablen. Hvis denne værdi (som tommelfingerregel) er mindre end 0,5, er sammenhængen så svag, at en identisk simulation med de samme tal kunne give en anden rangordning.

For inputvariabler, hvor der er en ikke-lineær sammenhæng mellem inputvariablen og outputvariablen, vil der være en tendens til, at regressionen indikerer en mindre sammenhæng mellem inputvariabel og outputvariabel, end der faktisk er. Der er i programmet ikke mulighed for at tage højde for ikke-lineære sammenhænge.

Tabel 6.2.9

Fejlanalyse af sammenhænge mellem inputvariabler og outputvariablen med multipel regression og Spearman rank korrelation.

Inputparameter	Følsomhedskoefficient	Rank korrelationskoefficient
Rug, udbytte (kg/ha)	-0.72	-0,75
Hørfrø Rusland, udbytte (kg/ha)	-0.20	-0.23
Rug, markarbejde (MJ/ha)	0.20	0.24
Hørfrø Rusland, anbefalet kg N (kg N/ha)	0.19	0.27
Rug, markarbejde (MJ/kg)	0.15	0.18
Hørfrø Rusland, markarbejde (MJ/ha)	0.15	0.15
Rug, pesticider (MJ/ha)	0.13	0.17
Rug (kg N/ha)	0.11	0.07
Hørfrø UK (kg N/ha)	0.11	0.11
Rug, anbefalet N (kg N/ha)	0.09	0.07

Analysen viser hvilke inputvariabler, der har størst indflydelse på variationen på outputvariablen. Indflydelsen på outputvariablen's variation vil være en funktion af, hvor meget inputvariablen har betydning for størrelsen af outputvariablen, og hvor stor usikkerheden er på inputvariablen.

Rangordenen i tabel 6.2.9 fortæller, at en reduktion af usikkerheden på bestemmelsen af udbyttet af rug pr. ha, kan have en betydelig indflydelse på usikkerheden på outputvariablen, hvorimod en bedre bestemmelse af anbefalet N pr. ha kun vil have lille indflydelse på usikkerheden på resultatet. Rangordningen kan bruges til at sætte en grænse for hvilke inputvariabler, som det næppe kan betale sig at bestemme mere nøjagtigt.

Det er muligt umiddelbart ved en simulering at vurdere, hvor stor indflydelse en bedre bestemmelse af en inputvariabel (eller flere variabler) vil have på usikkerheden på outputvariablen og på den baggrund at opstille en strategi for yderligere dataindsamling. Dvs. hvis målet f.eks. er at nedbringe standardafvigelsen på outputvariablen til det halve, kan man simulere, hvor meget usikkerheden på inputvariabler skal nedbringes. Der er i @Risk forskellige procedurer til at udføre følsomhedssimuleringer, hvor man på en enkel måde kan se, hvorledes ændringer i inputvariabler afspejles i outputvariablen.

En simulering, hvor det eksempelvis antages, at usikkerheden på udbyttet af rug pr. ha elimineres helt (standardafvigelse = 0), viser, at standardafvigelsen på outputvariablen ved dette reduceres fra 0,425 til 0,209. En eliminering af usikkerheden på bestemmelsen af forbruget af N til dyrkning af rug vil til sammenligning stort set ikke reducere standardafvigelsen på outputvariablen.

Cut off

I en livscyklusvurdering af f.eks. et produkt indgår der en række procesdata og en række systemdata, som kan være behæftet med usikkerheder, jf. tabel 4.2.1. Omfanget af dataindsamlingen vil ofte blive forsøgt reduceret ved bevidst fastsættelse af systemgrænserne og derved udelukkelse af visse led (dvs. *cut off*). *Cut off* kan foretages på flere niveauer, hvilket vil blive illustreret herunder.

Statistisk analyse kan i livscyklusvurdering anvendes til reducere omfanget af dataindsamlingen (*cut off*) ved:

- pege på led (produkter/processer), hvis bidrag til den samlede middelværdi er uvæsentligt, og som derfor kan udelades af den samlede livscyklusvurdering uden at påvirke resultatet nævneværdigt
- at pege på inputvariabler, hvor en reduktion af usikkerheden opnået ved yderligere dataindsamling ikke vil få væsentlig indflydelse på den samlede usikkerhed

1. metode

Det samlede energiforbrug til produktion af råvarer til 1 kg rugbrød er bestemt til 1,9 MJ med en standardafvigelse på 0,3 MJ. En metode til vurdering af *cut off* på et overordnet niveau vil være at relatere middelværdien for hver af de 8 processer opdelt på de 8 afgrøder til standardafvigelsen på det samlede energiforbrug. Middelværdierne er vist i tabel 6.2.10.

Tabel 6.2.10

Middelværdier for processer opdelt på de enkelte afgrøder.

	Rug	Sukkerrot	Hvede	Hørfrø,	Hørfrø,	Palmeolie	Sesamfrø	Solsikkefrø	Total
Pesticider	0,069	0,0033	0,01	0,0051	0,041	0,0008	0,0042	0,013	0,15
N-gødning	0,58	0,024	0,12	0,090	0,15	0,0094	0,022	0,059	1,04
P-gødning	0,033	0,0027	0,0060	0,0071	0,037	0,0003	0,0037	0,0057	0,095
K-gødning	0,014	0,0021	0,0025	0,0029	0,0068	0,0009	0,0010	0,0022	0,033
Kalk	0	0,0036	0,011	0	0	0	0	0	0,015
Husdyrgødning	0,0025	0,0001	0,0005	0	0	0	0	0	0,0031
Markarbejde	0,25	0,022	0,045	0,014	0,094	0,0031	0,0096	0,034	0,47
Tørring	0,12	0	0,026	0	0	0	0	0	0,15
Total	1,06	0,058	0,21	0,12	0,33	0,015	0,041	0,11	1,94

De fremhævede processer (angivet ved rastning) er ikke aktuelle for fremstilling af de pågældende afgrøder.

Ved at sammenholde totalerne for energiforbrug for henholdsvis processer og afgrøder med det samlede energiforbrug ses for processerne, at energiforbruget til fremstilling af P- og K-gødning, kalk og husdyrgødning ligger under standardafvigelsen eller for afgrøder, at energiforbruget til fremstilling af sukkerroer, palmeolie, sesamfrø og solsikkefrø ligger under standardafvigelsen. Tilsvarende kan gøres for hver enkelt afgrøde opdelt på processer.

2. metode

Ved fejlanalysen udført på baggrund af simuleringen i @RISK fås en rangordning af inputvariabler, hvor der ved fejlanalysen fokuseres på de inputvariabler, hvis usikkerhed giver det største bidrag til den samlede usikkerhed. I forbindelse med overvejelser om *cut off* fokuseres på de inputvariabler, der falder ud i den nederste ende af rangordningen, idet det er de inputvariabler, hvis usikkerhed ikke har nogen (væsentlig) indflydelse på den samlede usikkerhed. De lavt rangerede inputvariabler er der derfor ikke grund til at søge at bestemme bedre.

Sammenligning af to produkter

Til illustration af en livscyklusvurdering af 2 produkter, er der foretaget en sammenligning af energiforbruget til det beskrevne rugbrød med energiforbruget til et alternativt (muligt miljøforbedret) rugbrød, som udelukkende er baseret på danske råvarer.

I afsnit 6.2.1 er energiforbruget til produktion af 1 kg brød beskrevet. Dette brød er sammensat af såvel danske som udenlandske produkter. Til sammenligning benyttes en opskrift, hvor der udelukkende er benyttet danske produkter. Alle udenlandske afgrøder er her erstattet med dansk produceret rug.

Det estimerede energiforbrug samt variansen på energiforbruget til fremstilling af det danske rugbrød er vist i tabel 6.2.11. Tilsvarende oversigt for brødet, beskrevet i afsnit 6.2.1 ses i tabel 6.2.2.

Tabel 6.2.11

Estimeret energiforbrug til produktion af et rugbrød baseret på danske afgrøder. De udenlandske afgrøder er erstattet af dansk rug.

Afgrøde	Energiforbrug MJ/kg	Standardafvigelse MJ/kg	Varians (MJ/kg) ²	Variations-koefficient %
Rug	1,24970	0,30639	0,09388	25
Sukkerroer	0,05728	0,01159	0,00013	20
Hvede	0,21027	0,03311	0,00110	16
Totalt	1,51725	0,30839	0,09511	20

Det ses, at energiforbruget er lidt mindre for det danske brød, 1,52 MJ/kg, sammenlignet med den oprindelige opskrift med et energiforbrug på 1,94 MJ/kg. Usikkerheden på energiforbruget til brødet baseret udelukkende på danske afgrøder er 0,31 MJ/kg mod en usikkerhed på energiforbruget til det oprindelige brød på 0,30 MJ/kg med den oprindelige opskrift. Det ses her, at energiforbruget reduceres ved at

benytte udelukkende danske produkter, men det har ingen betydning for usikkerheden på energiforbruget.

I praksis vil man udelade de dele af livscyklus, der er ens for de to produkter og kun sammenligne de dele, der er ændrede fra det oprindelige produkt til det nye produkt. I nærværende eksempel skal energiforbruget til fremstilling af de udenlandske afgrøder således sammenlignes med energiforbruget til fremstilling af en tilsvarende mængde dansk rug. I tabel 6.2.12 er energiforbruget samt standardafvigelse og variationskoefficient vist for den del af de to brøds livscyklus, der er forskellig.

Tabel 6.2.12

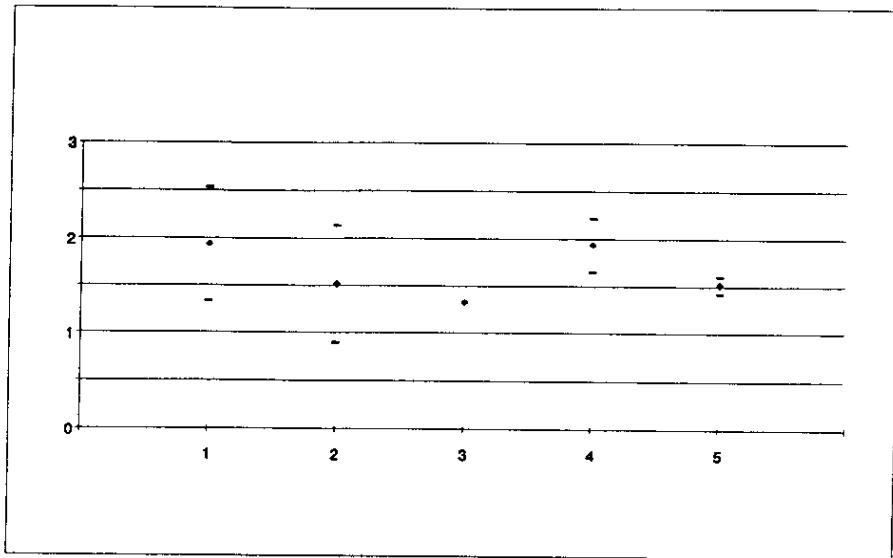
Estimeret energiforbrug til produktion af udenlandske afgrøder og en tilsvarende mængde dansk rug til et rugbrød samt standardafvigelse og varians.

Afgrøde	Energi (middelværdi) MJ/kg	Standard- afvigelse MJ/kg	Varians (MJ/kg) ²	Relativ varians %	Variations- koefficient %
Oprindelig opskrift					
Hørfrø, UK	0,11920	0,05401	0,0029	15	45,3
Hørfrø, Rusl.	0,32667	0,12221	0,0149	74	37,4
Palmeolie	0,01451	0,00448	0,0000	0,1	30,3
Sesamfrø	0,04057	0,01732	0,0003	1,5	42,7
Solsikkefrø	0,11322	0,04388	0,0019	9,6	38,7
Totalt	0,61417	0,14177	0,0201	100	23,1
"Dansk opskrift"					
Rug	0,18983	0,04654	0,0022	100	24,5

Statistisk analyse af forskellen

Det samlede energiforbrug til produktion af rugbrød efter den oprindelige opskrift og rugbrød med udelukkende danske afgrøder er bestemt til henholdsvis 1,94 MJ/kg og 1,52 MJ/kg. En test af om energiforbruget til produktion af rugbrødet baseret på danske afgrøder er signifikant mindre end energiforbruget med den oprindelige opskrift foretages et test af middelværdierne ved et t-test, som f.eks. er benyttet i Hoffmann *et al.* (1995). Approksimativt kan benyttes et test i normalfordelingen, idet der ikke haves kendskab til antal observationer og frihedsgrader. Benyttes en test i normalfordelingen ses ikke signifikant forskel ved anvendelse af et 5% signifikansniveau (p-værdien 0,33).

I figur 6.2.3 er energiforbruget til brød fremstillet efter den oprindelige opskrift og den alternative opskrift sammenlignet ved konfidensintervaller. Energiforbruget til fremstilling af de udenlandske afgrøder er ligeledes sammenlignet med energiforbruget til fremstilling af en tilsvarende mængde dansk rug. Figur 6.2.3 viser, at de samlede energiforbrug ikke er signifikant forskellige, men at de dele af livscyklus, der er forskellige (den del der er "forbedret"), er signifikant forskellige.



Figur 6.2.3

Sammenligning af energiforbruget til produktion af 1 kg brød med to forskellige opskrifter ved anvendelse af konfidensintervaller. 1 og 2 viser det totale energiforbrug ved fremstilling af det oprindelige og det alternative rugbrød. 3 viser energiforbruget til den fælles del af de to rugbrød mens 4 og 5 viser det supplerende energiforbrug for fremstilling af udenlandske afgrøder henholdsvis dansk rug.

6.3 Case 2: Emissioner fra bageri

6.3.1 Statistisk analyse af data

I nærværende afsnit beskrives den statistiske analyse foretaget af Case 2: emission fra bageri. Selve casen er beskrevet i afsnit 5.3.1. Data og datakvalitet er beskrevet i afsnit 5.3.2. Disse 2 afsnit er udgangspunktet for den nedenfor beskrevne statistiske analyse.

Til beregning af udledning af BOD og total-N indgår følgende processer mht. udledningen:

- Produktion af rugbrød
- Vandforbrug i produktionen af brød
- Spildevand
- Udledning af BOD pr. brød
- Udledning af total-N pr. brød

Produktion af rugbrød

Produktionen af rugbrød kendes pr. måned for 1992 svarende til 21.738 ton brød/år. Dette er gennemsnitligt 1811,5 ton pr. måned med en standardafvigelse på 165,3 ton/måned. Standardafvigelsen på den årlige produktion er bestemt ud fra standardafvigelsen på den gennemsnitlige produktion pr. måned:

$$std(brød\ i alt) = \sqrt{12} \cdot 165,3 = 572,62 \text{ ton brød/år} \quad (50)$$

Vandsforbrug i produktionen af brød

Vandsforbruget på bageriet er $613,67 \text{ m}^3$ vand/uge med en usikkerhed på $91,39 \text{ m}^3$ /uge.

Vandsforbruget pr. år på bageriet og standardafvigelsen beregnes, jf. (formel xx), til

$$\begin{aligned} E(vand\ i\ alt) &= 52 \text{ uger} \cdot 613,67 \text{ m}^3/\text{uge} = 31.910 \text{ m}^3/\text{år} \\ std(vand\ i\ alt) &= \sqrt{52} \cdot 91,39 \text{ m}^3/\text{år} \end{aligned} \quad (51)$$

Vandsforbruget pr. kg brød og standardafvigelsen beregnes, jf. (formel xx), ved

$$E(vand\ pr.\ brød) = \frac{31.910 \text{ m}^3 \text{ vand/år}}{21.738 \text{ ton brød/år}} = 1,468 \text{ m}^3 / \text{ton brød} \quad (52)$$

$$\begin{aligned} Var(vand\ pr.\ brød) &= \frac{1}{(21.738 \text{ ton brød/år})^2} (658 \text{ m}^3 \text{ vand/år})^2 \\ &+ \frac{(31.910 \text{ m}^3 \text{ vand/år})^2}{(2.738 \text{ ton brød/år})^4} (572,62 \text{ ton brød/år})^2 \\ &= 0,0024 \text{ (kg vand/kg brød)}^2 \end{aligned} \quad (53)$$

$$std(vand\ pr.\ brød) = \sqrt{0,0024 \text{ kg (vand/kg brød)}^2} = 0,0491 \text{ kg vand/kg brød} \quad (54)$$

som vist i tabel 6.3.1.

Tabel 6.3.1

Vandsforbrug til fremstilling af 1 kg brød.

Vandsforbrug pr. brød	Estimat kg vand/kg brød	Standardafvigelse kg vand/kg brød
	1,468	0,0491

Spildevand

Vandsforbruget er beregnet til $1,468 \text{ kg vand/kg brød}$. Der tilskættes i alt $0,33 \text{ kg vand/kg brød}$ med en anslået variationskoefficient på 1% (standardafvigelse på $0,0033 \text{ kg vand/kg brød}$).

Den udledte mængde spildevand og standardafvigelse, jf. (formel xx), estimeres herved til

$$E(spv.) = 1,468 - 0,33 = 1,138 \text{ kg vand/kg brød} \quad (55)$$

$$std(spv.) = \sqrt{0,0024146 + 0,000011} = 0,049 \text{ kg vand/kg brød} \quad (56)$$

Udledning af BOD og total-N pr. brød

Udledningen af BOD og total-N foregår af 2 udledninger, her angivet Nord (N) og Øst (Ø), som angivet i tabel 6.3.2.

Tabel 6.3.2

Udledning af BOD og total-N.

Udledning	Sted	Estimat mg/l	Standardafvigelse mg/l
BOD	nord	2.201,90	1300,0
BOD	øst	78,45	68,1
Total-N	nord	66,31	62,7
Total-N	øst	4,69	7,3

Forholdet mellem mængden af udledt vand i de 2 udledninger er ikke kendt. Beregningen af usikkerheden på udledningen af BOD og total-N er vist ved to forskellige metoder:

- I: Der antages dels forholdet 1:1, dels "worst case", 9:1, hvor det antages, at 90% af udledningen sker af det nordlige udløb, hvor der er fundet de højeste koncentrationer af BOD og total-N.
- II: Der regnes med middelværdien for forholdet 1:1 med en samlet usikkerhed baseret på standardafvigelsen på de enkelte målinger korrigteret for usikkerheden ved, at det ikke vides, hvordan udledningen fordeler sig mellem det østlige og nordlige udløb, jf. tabel 5.3.8.

Metode I:

Beregning af hhv. BOD og total-N pr. kg brød bliver herved

$$E(\text{udledning}) = (\text{ratio}_N X_N + \text{ratio}_\Theta X_\Theta) Y \quad (57)$$

hvor ratio_N og ratio_Θ er andelen af spildevand udledt i hhv. Nord og Øst (1:1 eller 9:1)

X_N og X_Θ er mængden af BOD (eller total-N) i hhv. nordlig og østlig udledning

Y er beregnet mængde spildevand ved produktion af et brød

Herved fås følgende resultater for emission af BOD og total-N, som angivet i tabel 6.3.3.

Tabel 6.3.3

Udledning af BOD og total-N pr. kg brød samt standardafvigelser beregnet ved metode I.

Udledning	Forhold Nord-Øst	Estimat	Standardafvigelse	Variations- koeficient %
		mg/kg brød	mg/kg	
BOD	1:1	1.29752	0.05794	4,46
BOD	9:1	2.26411	0.10389	4,58
Total-N	1:1	0,040393	0.00503	12,41
Total-N	9:1	0,068438	0.00863	12,57

"Worst case" betragtningen viser, at for BOD sker der en fordobling i udledningen ved ændring af ratioen mellem nord og øst fra 1:1 til 9:1, og for total-N bliver den tilsvarende stigning ca. 40%. I begge tilfælde vil det derfor være relevant at søge mere præcise oplysninger om ratioen mellem spildevand udledt i henholdsvis nord og øst.

Metode II

Beregning af BOD og total-N pr. kg rugbrød bliver hermed

$$\begin{aligned} E(\text{udledning}) &= (\text{ratio}_N X_N + \text{ratio}_\Theta X_\Theta) Y \\ &= (X_N + X_\Theta) \text{ratio } Y \end{aligned} \quad (58)$$

hvor ratio_N og ratio_Θ er andelen af spildevand udledt i hhv. Nord og Øst og $\text{ratio} = 0,5$, idet der her er antaget forholdet 1:1 i udledningen mellem Nord og Øst

X_N og X_Θ er mængden af BOD (eller total-N) i hhv. nordlig og østlig udledning

Y er beregnet mængde spildevand ved produktion af et brød

Standardafvigelsen bestemmes ved (12) med en varianskoeficient for ratioen givet i tabel 5.3.8 til henholdsvis 37,3% og 37,2% for BOD og total-N, se tabel 6.3.4.

Tabel 6.3.4

Udledning af BOD og total-N pr. kg brød samt standardafvigelser beregnet ved metode II.

Udledning	Forhold Nord-Øst	Estimat	Standardafvigelse	Variations- koeficient %
		mg/kg brød	mg/kg	
BOD	1:1	1.29752	0,48357	37,26
Total-N	1:1	0,040393	0,01501	37,22

6.3.2 Vurdering af resultaterne af den statistiske analyse af data

Emissionen af BOD og total-N pr. kg brød er estimeret til hhv. 1,30 og 0,04 mg/kg. Det er her antaget, at forholdet mellem vandmængderne i nordlig og østligt udløb er ens. Usikkerhederne til disse estimater er relative små (svarende til en CV på hhv. 4,5% og 12,5%). Hvis mængden af vand, udledt i nordlig retning derimod er væsentligt større end i østlig retning, får dette en stor indflydelse på estimatet af emissionen af BOD og total-N. Her er benyttet, at 90% af vandet udledes i nordlig retning, hvorved udledningen af BOD og total-N er henholdsvis 2,26 og 0,068 mg/kg brød (metode I).

Den potentielle fejl, der introduceres ved at fordelingen af vand i det nordlige og det østlige ikke er kendt, kan ligeledes afspejles i den supplerende usikkerhed på det gennemsnitlige indhold af BOD og total-N i udløbet. Emissionen af BOD og total-N pr. kg brød er igen estimeret til hhv. 1,30 og 0,04 mg/kg, men variationskoefficienten på estimatet bliver nu 37% (metode II).

Hvorvidt middelværdien bør korrigeres ud fra en "worst case" betragtning, eller den øgede variationskoefficient bør anvendes til at give et retvisende billede kan ikke afgøres ud fra dette eksempel alene.

6.4 Case 3: Levetid for køleskabe

I nærværende afsnit er det årlige energiforbrug til køleskabe beregnet med speciel opmærksomhed på betydningen af brug af forskellige fordelinger af et køleskabs levetid.

6.4.1 Statistisk analyse af data

Beskrivelse af energiforbruget til nedkøling af varer er beskrevet i afsnit 5.4. Køleskabenes reelle levetid kendes ikke (jf. afsnit 5.4). Derimod kendes salgstallet for køleskabe. Disse oplysninger er benyttet til estimation af antal køleskabe i funktion som alternativ til den reelle levetid.

På basis af salgstallene og en antagelse om et mættet marked er middellevetiden bestemt til 14 år, jf. 5.4.1. Usikkerheden er vurderet til en variationskoefficient på 25%, hvilket giver en standardafvigelse på 3,5 år.

Der er i det følgende foretaget beregning af levetiden samt energiforbruget på basis af dels en normalsfordeling, dels en gammafordeling. Da de reelle levetider er ukendte, er det ikke muligt at teste og vurdere, hvorvidt de forskellige fordelinger er korrekte. Eksemplet kan derimod benyttes til at illustrere, hvilke forskelle der fremstår ved anvendelse af forskellige fordelinger og hermed betydningen af valg af den rigtige fordeling i andre eller lignende situationer.

Antages det, at levetiden X følger en normalsfordeling, haves

$$X \in N(\mu, \sigma^2) \quad (59)$$

hvor μ og σ^2 er hhv. middelværdi og varians i fordelingen.

Estimat af disse 2 parametre er hhv. 14 år og $(3,5 \text{ år})^2$, og levetiden følger hermed fordelingen

$$X \in N(14; 3,5^2) \quad (60)$$

Antages derimod en gammafordeling, haves

$$X \in G(k, \beta) \quad (61)$$

hvor β er en skalaparameter og k en formparameter med middelværdi og varians:

$$E(X) = k \beta \quad (62)$$

$$V(X) = k \beta^2 \quad (63)$$

Dette kan benyttes ved estimation af fordelingens 2 parametre, k og β . Approximativt anvendes estimerer for k og β på hhv. 16 og 0,875. Herved følger levetiden fordelingen

$$X \in G(16; 0,875) \quad (64)$$

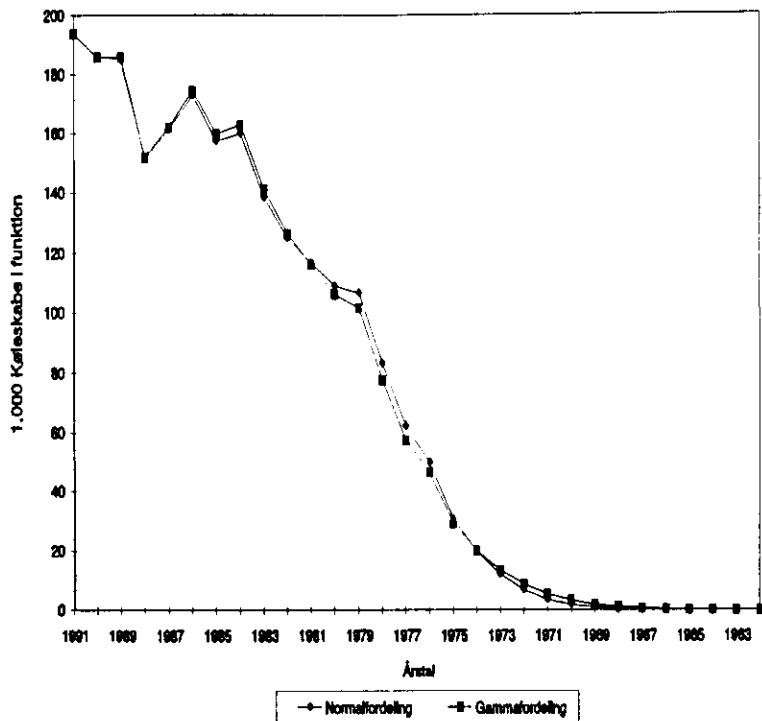
Antallet af køleskabe i funktion i 1992 er bestemt ved antagelse om forskellig fordeling af levetiden. En normalfordeling og en gammafordeling med parametre som vist ovenfor.

Tabel 6.4.1

Sammenligning af forskellige fordelinger til beskrivelse af levetiden for køleskabe. Levetiden er beskrevet enten ved en normalfordeling $X \in N(\mu, \sigma^2)$ eller ved en gammafordeling $X \in G(k, \beta)$.

Salgsår (alder)	Solgte køleskabe 1.000 stk	Antal køleskabe i funktion 1.000 stk	
		N(14;3.5 ²)	G(16;0.875)
1991 (1)	194	193.980	194.000
1990 (2)	186	185.944	186.000
1989 (3)	186	185.844	186.000
1988 (4)	152	151.675	151.996
1987 (5)	162	161.180	161.951
1986 (6)	175	173.051	174.655
1985 (7)	161	157.337	169.675
1984 (8)	167	159.779	162.834
1983 (9)	150	138.515	141.093
1982 (10)	143	124.903	126.233
1981 (11)	145	116.626	116.030
1980 (12)	152	108.854	105.991
1979 (13)	174	106.567	101.393
1978 (14)	166	83.000	77.480
1977 (15)	160	62.008	57.379
1976 (16)	175	49.675	46.330
1975 (17)	155	30.331	29.167
1974 (18)	155	19.615	20.005
1973 (19)	155	11.867	13.271
1972 (20)	155	6.702	8.535
1971 (21)	155	3.526	5.332
1970 (22)	155	1.726	3.243
1969 (23)	155	0.785	1.923
1968 (24)	155	0.331	1.114
1967 (25)	155	0.130	0.652
1966 (26)	155	0.047	0.351
1965 (27)	155	0.016	0.191
1964 (28)	155	0.005	0.102
1963 (29)	155	0.001	0.054
1962 (30)	155	0.000	0.028
I alt i funktion i 1992		2.234.021	2.232.985

Fordelingerne er ligeledes vist i figur 6.4.1.



Figur 6.4.1

Sammenligning af forskellige fordelinger til beskrivelse af levetiden for køleskabe. Levetiden er beskrevet ved enten en normalsfordeling $X \in N(\mu; \sigma^2)$ med $(\mu; \sigma^2) = (14,3; 5^2)$ eller en gammafordeling $X \in G(k, \beta)$ med $G(k, \beta) = (16; 0,875)$.

Energiforbruget i 1992 er beregnet for køleskabe fra 1970-1991 ved anvendelse af de 2 forskellige fordelinger for antal køleskabe i funktion. Energiforbruget med angivelse af usikkerhed og varianskoefficient ses for de 2 fordelinger i tabel 6.4.2.

Tabel 6.4.2

Energiforbrug (MWh) for køleskabe i 1992 ved anvendelse af forskellige antagelser om fordeling af antal køleskabe i funktion i 1992.

Fordeling	Energiforbrug MWh	Standardafvigelse MWh	Variationskoefficient %
$N(14,3; 5^2)$	770.531	79,01	0,0103
$G(16; 0,875)$	770.058	79,02	0,0103

6.4.2 Vurdering af resultaterne af den statistiske analyse af data
Antallet af køleskabe i funktion er beregnet ved anvendelse af en normal- og gammafordeling. Parameterværdierne for gammafordelingen er valgt således, at den gennemsnitlige levetid for et køleskab er uændret. Det ses, at antallet af køleskabe i funktion er meget ens for de to fordelinger.

Energiforbruget for køleskabe i 1992 ses at variere meget lidt afhængigt af antaget fordeling for antal køleskabe i funktion. Usikkerheden

på energiforbruget er tilsvarende lille for de to fordelinger. I nærværende eksempel ses effekten af valg af fordeling for antal køleskabe i funktion at være af meget lille betydning for usikkerheden på køleskabs energiforbrug.

6.5 Sammenfattende vurdering af den statistiske analyse

En metode til statistisk analyse af data i en livscyklusvurdering er introduceret og beskrevet i afsnit 6.1.1.

Den statistiske analyse består i principippet af følgende tre trin:

- Opstilling af procesforløbet, som skal benyttes i livscyklusvurderingen
- Antagelse om fordeling hvis der anvendes simuleringer, samt parameterværdierne i disse fordelinger, f.eks. normalfordeling med middelværdi og varians
- Beregning af middelværdi og varians for den samlede proces på basis af procesforløbet og de antagte fordelinger og parameterværdier

Herved opnås et estimat af middelværdi og varians (usikkerhed, CV, etc.) for processen.

Disse beregninger er i nærværende rapport bestemt for en enkelt case ved anvendelse af approksimative beregninger uden antagelse om inputvariablens fordelinger samt ved simulering med antagelser om samtlige inputvariablers fordelinger. De approksimative beregninger er enkle at anvende, så længe procesforløbet ikke er for kompliceret. For store og komplicerede procesforløb kan simulering anvendes for at reducere store approksimative beregninger.

Simulering bygger på specifikke fordelingsantagelser, som har betydning for resultatet af simuleringen. Analyserne viser, at fordelingen af outputvariablen ikke er symmetrisk. Middelværdien ligger over værdien fundet ved "traditionel" beregning af outputvariablen. Ved traditionel beregning forstås den metodik, der normalt benyttes ved f.eks. beregning af det samlede energiforbrug for et produkts livscyklus, hvor der direkte regnes med middelværdier.

I de 2 sidste cases er udelukkende anvendt appproximative beregninger.

Efter beregning af middelværdi og varians for procesforløbet, kan der være behov for en række yderligere beregninger til f.eks.:

- sammenligning af 2 eller flere parallelle produkter
- fejlanalyse
- vurdering af *cut off*

For den 1. case: "Energiforbrug til produktion af 1 kg rugbrød" er disse beregninger illustreret. Dette har vist, at der kan være stor

forskelt på, hvilke led i det samlede procesforløb, som bidrager væsentligt til den samlede varians.

Dette har også betydning for muligheden for *cut off*, hvor der tænkes på reduktion af procesforløbet til et mere enkelt forløb. Der er skitseret to metoder til at foretage *cut off*, hvor den første metode er baseret på de inputvariabler, der rangerer lavet i fejlanalysen og den anden metode er baseret på middelværdier relateret til standardafvigelsen på det samlede resultat. Ved den sidstnævnte metode er det ikke umiddelbart muligt at relatere de "afskårne" middelværdier til inputvariablerne.

Sammenligning af 2 eller flere produkter kan enkelt foretages ved at sammenligne konfidensintervaller for de fundne middelværdier. En anvendelse af livscyklusvurdering er at vurdere et forbedret produkt i forhold til det oprindelige. Hvis det oprindelige og det nye produkt har dele af livscyklus fælles, sammenlignes kun de dele, der er forskellige. Det vil i så fald ofte være muligt at vise signifikante forskelle, som ikke ville fremgå ved sammenligning af produkternes samlede livscyklus.

7. Internationalt seminar om statistik og usikkerheder i forbindelse med livscyklusvurderinger

Som led i projektet inviterede projektgruppen til et "Invitational Expert Seminar on Uncertainty and Statistics for Product Life Cycle Assessments" i Delft d. 23. maj 1995. Det praktiske arrangement af seminaret blev forestået af Maarten van Hees, Ph.D.-studerende ved Delfts Tekniske Universitet, og udgifterne blev dækket af de enkelte deltagere. Projektgruppen inviterede en række eksperter inden for henholdsvis livscyklusvurdering (primært arbejdsgruppen om data håndtering under SETAC (Society for Environmental Toxicology and Chemistry)) og statistisk behandling af komplekse systemer. I alt 14 eksperter deltog, hvoraf 6 havde statistisk baggrund, og resten erfaring i livscyklusvurdering.

Der blev givet 4 mundtlige oplæg, hvoraf det ene var projektgruppens præsentation af en tidlig version af kapitel 5.1. af nærværende rapport. Ét indlæg fokuserede meget på proces-specifik statistik. Det ene af de sidste to indlæg (Heijungs 1995) er refereret i kapitel 4.

Efter de 4 indlæg blev deltagerne delt op i 2 arbejdsgrupper med emnerne:

- 1) Vurdering af usikkerhed på grund-data
- 2) Strategi til vurdering af pålidelighed af resultater af livscyklusvurderinger

Arbejdsgruppe I

Den første arbejdsgruppe beskæftigede sig primært med emnet for nærværende rapport på baggrund af projektgruppens oplæg. Arbejdsgruppen kom frem til følgende anbefalinger:

- En datakvalitets-målsætning bør altid være en del af en livscyklusvurderings målformulering og afgrænsning.
- Der bør altid angives en usikkerhed, også når den faktiske usikkerhed er ukendt. I så fald skal usikkerheden skønnes på baggrund af sammenligninger med analoge situationer, hvor usikkerheden er kendt eller ved hjælp af standard-antagelser (jf. kapitel 5.1.5.)
- Når den faktiske fordeling er ukendt, og der ikke er grund til at antage andet, bør data antages at være normalfordelt. Som begrundelse for denne anbefaling blev det fremført, at den endelige fordeling under alle omstændigheder vil nærme sig en normalfordeling, når et stort antal uafhængige variabler lægges sammen. Imidlertid blev der i det efterfølgende plenum stillet spørgsmålstege ved, om denne anbefaling var rimelig og, om den overhovedet var nødvendig.

- I databaser og publicerede livscyklus-studier bør usikkerhed angives, helst ved brug af videnskabelig notation for data og usikkerhed udtrykt f.eks. ved standardafvigelsen. Som et minimum bør der angives det korrekte antal betydende cifre.
- Der er behov for en ordliste/konvention til at sikre en ensartet brug af begreber som pålidelighed, præcision (precision/accuracy), følsomhed, fejl, usikkerhed, \pm , etc.
- Der er behov for at fremskaffe flere eksempler på usikkerheder, baseret på faktiske beregnede værdier for basale råvarer og processer.
- Grundlæggende er der brug for mere erfaring med brug af usikkerhed i livscyklusvurdering i praksis.

Arbejdsgruppe II

Den anden arbejdsgruppe udarbejdede en strategi til vurdering af pålideligheden af resultater af livscyklusvurderinger. Udgangspunktet var, at mange livscyklusvurderinger publiceres uden angivelse af usikkerheder på resultaterne (endsige på de tilgrundliggende data). Når man ikke har adgang til de oprindelige data i elektronisk form, bliver det en uoverskuelig opgave at beregne den samlede usikkerhed, selv i det tilfælde hvor man kender eller kan estimere usikkerheden på de grundliggende data. Der er derfor behov for en strategi til på simpel måde at estimere usikkerheden på et resultat af en publiceret livscyklusvurdering.

Resultatet af livscyklusvurderinger foreligger ofte i form af et eller flere tal for hvert af de sammenlignede produktsystemer. Disse tal kan være fremkommet ved sammenvejninger af data fra en lang række enkelprocesser og vægtningsfaktorer. Arbejdsgruppen foreslog følgende procedure til vurdering af usikkerheden på slutresultatet:

1. Først udvælges de 3 største bidrag til det samlede resultat.
2. Dernæst udvælges de 3 største bidrag til disse og så fremdeles, indtil man når ned til de bidragende enkelprocesser. På dette niveau estimeres det største bidrag til usikkerheden. Denne usikkerhed antages at være et dækkende estimat for usikkerheden på alle de indgående data.
3. Her ud fra estimeres usikkerheden på det næste niveau op og så fremdeles, indtil man atter når det samlede resultat.

I tilfælde af produktsammenligninger er det de største bidrag til forskellen mellem de analyserede produktsystemer, der skal følges.

Den usikkerhed, der ved denne metode estimeres for det samlede resultat, er sandsynligvis lidt lavere end den faktiske usikkerhed. De opnåede estimer vil evt. kunne korrigeres ved hjælp af erfaringsbaserede korrektionsfaktorer, der kan etableres ved at afprøve metoden på en række livscyklusvurderinger, hvor den faktiske usikkerhed kendes i forvejen. En sådan afprøvning vil endvidere kunne lede til en korrektion i selve metoden, idet det er muligt, at der skal udvælges et andet antal bidrag end 3 for at få et rimeligt estimat. Det kan være, at det optimale antal er 5, eller at man bør udvælge bidrag svarende til en bestemt procentdel af de indgående bidrag.

Arbejdsgruppen anbefalede i øvrigt, at diskrete usikkerheder, som f.eks. ved valg mellem flere mulige allokeringsmetoder, behandles ved simulering af de enkelte alternativers indflydelse på resultatet.

Resultatet fra seminaret vil indgå i det fremtidige arbejde med en forbedret "Code-of-Practice" for livscyklusvurderinger fra Society for Environmental Toxicology and Chemistry (SETAC).

8. Anbefalinger til videre arbejde/projektoplæg

Gennemførelse af nærværende projekt har vist, at der er en række punkter, hvor metoden til beskrivelse af datakvalitet (afsnit 5.1) kræver afprøvning i en bredere kreds af brugere med henblik på at teste om forskellige brugere når frem til samme resultat. De statistiske beregninger gennemført i de udvalgte cases peger ligeledes på, at der er behov for at arbejde videre med test og vurdering af de forskellige metoder herunder afprøvning af kommercielle programmer til livscyklusvurdering, som indeholder muligheder for behandling af usikkerheder på data. I de præsenterede casestudier er f.eks. datakvaliteten beskrevet meget detaljeret med henblik på at vise, hvordan metoden kan anvendes. I et enkelt tilfælde er der givet et eksempel på forenkling af den sammenfattende tabel med angivelse af datakvaliteten for et datasæt, se tabel 5.3.9 og 5.3.10. Der vil være behov for en mere uddybende diskussion af krav til beskrivelse af datakvalitet i forskellige situationer.

Forslag til videre arbejde er herunder nævnt på punktform, og de er uddybende beskrevet i det efterfølgende.

- Afprøvning af metoden til beskrivelse af datakvalitet i en gruppe af brugere - scoring af datakvalitetsindikatorer og skøn/bestemmelse af supplerende usikkerheder betinget af datakvaliteten med henblik på at forbedre metoden og skabe consensus om den praktiske anvendelse mv.
- Afprøvning af metoden til beskrivelse af datakvalitet på en række cases udvalgt til belysning af forskellige aspekter.
- Udarbejdelse af paradigma for beskrivelse af datakvalitet på forskellige niveauer.
- Test af kommersielt software til statistisk behandling af usikkerheder i forbindelse med livscyklusvurdering sammen med yderligere afprøvning af såvel @RISK som SAS.
- Indsamling af industrielle data inden for udvalgte sektorer til brug for bestemmelse af generelle usikkerheder på industrielle data.
- Usikkerheder på vurderingsdata forekommer, og betydningen heraf bør undersøges nærmere.
- Korrelationer mellem inputvariabler kan påvirke outputvariablen, hvis ikke størrelsen af korrelationerne kendes og medtages i beregningerne.

- *Cut-off* og fejlanalyse er kort berørt, men der er behov for yderligere beskrivelse af og test af metoderne med henblik på at opnå konsensus om, hvordan metoderne kan anvendes i livscyklusvurdering.
- Generel vejledning i anvendelse af usikkerhedsvurderinger i livscyklusvurdering. Vejledningen skal beskrive forudsætninger, metoder samt dokumentere de foreliggende metoder.

Metodeafprøvning I

Under arbejdet med beskrivelse af datakvaliteten har det vist sig, at de anvendte datakvalitetsindikatorer ikke i alle tilfælde entydigt beskriver data i de bearbejdede cases, ligesom tildelingen af en score til de forskellige indikatorer ikke i alle tilfælde er entydig. Asprøvning af metoden i en bredere gruppe af livscykluspraktikere vil kunne give grundlag for forbedring og validering af metoden. En sådan asprøvning vil kunne sige noget om reproducerbarheden af beskrivelsen af datakvaliteten. Afhængig af udfaldet af asprøvningen vil resultaterne kunne bruges til forbedring af de forskellige datakvalitetsindikatorer såvel som forbedring af beskrivelsen af scoringen af de forskellige datakvalitetsindikatorer.

Metodeasprøvningen vil ligeledes kunne sige noget om reproducerbarheden af skøn/bestemmelse af supplerende usikkerheder betinget af datakvaliteten. Yderligere præcisering heraf kræver en mere detaljeret indsamling af industrielle data inden for udvalgte sektorer (se senere).

Metodeafprøvning II

Metoden til beskrivelse af datakvalitet er indtil videre kun asprøvet i begrænset omfang - i nærværende projekt på tre cases. Asprøvningen af metoden på flere cases f.eks. i forbindelse med asprøvning af metodestandarder (ISO, NM) vil sammen med en asprøvning af metoden i en gruppe af uafhængige livscykluspraktikere kunne bidrage til en yderligere forbedring af metoden.

Paradigma for beskrivelse af datakvalitet

Under forudsætning af at der opnås en form for konsensus om metoden, bør der udarbejdes et paradigme for beskrivelse af datakvalitet. Metoden bør udbredes til andre områder end livscyklusvurderinger. Det drejer sig om beskrivelse af data i andre sammenhænge f.eks. andre miljøvurderinger, VVM, grønne regnskaber, EMAS såvel som offentlige og private virksomheders interne og eksterne dokumentation af datakvalitet. Som tidligere nævnt kan den systematiske anvendelse af metoden til beskrivelse af datakvaliteten ud fra datakvalitetsmatricen betragtes som en tjekliste. Rapportering af datakvaliteten under hensyntagen til en given datakvalitetsmålsætning vil rejse et behov for en udvikling af et paradigma. Paradigmaet skal medvirke til at sikre en ensartet og tilstrækkelig beskrivelse af data.

Test af software

Kommerciel tilgængeligt software til anvendelse i forbindelse med livscyklusvurdering indeholder evt. muligheder for at inddrage usikkerheder i beregningerne. Det drejer sig f.eks. om KCL-ECO (udviklet af The Finnish Pulp and Paper Research Institute).

Udvalgte programmer anskaffes, og der udarbejdes en beskrivelse og analysemetode, hvormed usikkerheder behandles. Programmer testes på udvalgte cases. @RISK og SAS er hver for sig blevet afprøvet af øvede brugere, hvorfor det vil være relevant at lade uøvede brugere afprøve samme metoder med henblik på at vurdere potentialet for anvendelse af de forskellige metoder. Usikkerheder behandles samtidig med eksakte metoder ved programmering i SAS og/eller ved simuleringer i @RISK såvel som SAS.

Industrielle data

Der udvælges en række sektorer og inden for hver sektor en række virksomheder, som inddrages i et projekt som leverandører af data. Det kan evt. komme på tale at anvende data indsamlet i andre projekter f.eks. i forbindelse med generel dataindsamling og databaseopbygning.

Vurderingsdata

Vurderingsdata vil ofte være behæftet med usikkerheder. Der er behov for en systematisk gennemgang af de vurderingsdata, der anvendes i livscyklusvurdering, med henblik på at beskrive og kvantificere usikkerhederne.

Korrelationer

I nærværende projekt er det antaget, at samtlige inputvariabler er ukorrelerede. Det er imidlertid ikke korrekt, hvorfor antagelsen indfører en fejl i det endelige resultat. Betydningen af korrelationerne afhænger af størrelsen af disse.

Der er flere problemer i relation til relationer mellem inputvariabler. Her skal først og fremmest nævnes, at det kan være vanskeligt at skaffe information om størrelsen af disse korrelationer. Endvidere kan det være vanskeligt at inddarbejde korrelationernes betydning i en livscyklusvurdering. Der er behov for at arbejde med cases, hvor der primært fokuseres på at identificere korrelationer og vurdere betydningen heraf.

Cut-off/fejlanalyse/følsomhedsanalyse

Der er ikke ved litteratursøgningerne fundet litteratur, som præcist definerer de tre begreber i forhold til livscyklusvurdering. I nærværende projekt er fejlanalyse anvendt, og der er givet nogle bud på, hvordan *cut-off* begrebet kan håndteres. Der er derfor behov for at arbejde yderligere med definition af de tre begreber og afprøve dem på udvalgte cases. Afprøvningen af metoderne på cases skal lede til en operationalisering af metoderne i forhold til livscyklusvurdering.

Vejledning i statistisk analyse i livscyklusvurdering

Flere af ovennævnte punkter leder frem til, at der udarbejdes en decideret vejledning i anvendelse af statistisk analyse i livscyklusvurdering. Vejledningen vil kunne indeholde:

- statistiske metoder i relation til livscyklusvurdering
- beskrivelse af software til håndtering af statistiske problemer i forbindelse med livscyklusvurderinger
- metoder til håndtering af korrelationer mellem inputvariabler

- metoder til *cut-off*, fejlanalyse og følsomhedsanalyse
- dokumentation af de foreslæde metoder

Referencer

Bomholt JK, Pedersen PN, Danielsen V, Nielsen NO, Nørgaard E, Udesen FK, Bay HP (1992). Håndbog for svinehold 1993. Århus: Landbrugets Informationskontor.

Borland International (1992). Quattro® Pro. User's guide, version 4.0. Scotts Valley, CA: Borland International Inc.

Boustead I (1992). Eco-balance methodology for commodity thermoplastics. Brussels: APME/PWMI.

Boustead I (1993a). Eco-profiles of the European plastics industry. Report 2: Olefin feedstock sources. Brussels: APME/ PWMI.

Boustead I (1993b). Eco-profiles of the European plastics industry. Report 3: Polyethylene and polypropylene. Brussels: APME/PWMI.

Boustead I (1993c). Eco-profiles of the European plastics industry. Report 4: Polystyrene. Brussels: APME/PWMI.

Boustead I (1994). Eco-profiles of the European plastics industry. Report 6: Polyvinyl chloride. Brussels: APME/PWMI.

Conradsen K (1984). En introduktion til statistik. Bind 1A og 1B. 5. udgave 1984. Lyngby: IMSOR, DTH.

Constantinou E, Seigneur C (1991). Development of multimedia health risk assessment model. Proc. HAZMACON 91, Santa Clara, California, April 16, 1991.

Constantinou E, Seigneur C, Permutt T (1992). Uncertainty analysis of health risk estimates - Application to power plant emissions. For presentation at the 85th Annual Meeting & Exhibition, Kansas City, Missouri June 21-26, 1992.

Franklin Associates LtD (1993). Statistical considerations of data error and variability. (Unpublished)

Funtowicz SO, Ravetz JR (1994). Uncertainty and quality in science for policy. Dordrecht: Klüwer Academic Publishers.

Green MB (1976). Energy in agriculture. Chemistry and Industry 641-646.

Green MB (1987). Energy in pesticide manufacture, distribution and use. pp. 165-177 in Helsel ZR (ed.): Energy in plant nutrition and pest control. Amsterdam: Elsevier (Energy in World Agriculture no. 2).

Green MB, McCulloch A(1976). Energy considerations in the use of herbicides. *J. Sci. Fd. Agric.* 27:95-100.

Heijungs R (1994). A generic method for the identification of options for cleaner products. *Ecological Economics* 10:69-81.

Heijungs R (1995). On the identification of key issues for further investigation in life-cycle screening. The use of mathematical tools and statistics for sensitivity analyses. Presented to the Invitational Expert Seminar on Uncertainty and Statistics for Product Life Cycle Assessment, Technical University of Delft, 1995.05.23.

Heijungs R, Guinée JB, Huppes G, Lankreijer RM, Udo de Haes HA, Sleeswijk AW, Ansems AMM, Eggels PG, van Duin R, de Goede HP (1992). Environmental life cycle assessment of products. I Guide; II Backgrounds, 1992. Leiden: CML.

Hoffmann L, Weidema BP, Christiansen K, Ersbøll AK (1995). Statistical analysis and uncertainties in relation to LCA. In Lindfors L-G (ed.). *LCA-Nordic - Technical Reports No 10 and Special Reports No 1 - 2*. TemaNord 1995:503. Copenhagen: Nordic Council of Ministers.

ISO (1993). Statistics - Vocabulary and symbols - Part 1: Probability and general statistical terms. Reference number ISO 3534-1:1993 (E/F).

Lindfors L-G, Christiansen K, Hoffmann L, Virtanen Y, Juntilla V, Hanssen OJ, Rønning A, Ekwall T, Finnveden G (1995). *Nordic Guidelines on Life-Cycle Assessment*. Nord 1995:20. Copenhagen: Nordic Council of Ministers.

Microsoft (1994). Microsoft Excel, Version 5.0c. Microsoft Corporation.

Palisade Corporation (1991). Risk analysis and simulation add-in for Microsoft Excel - Windows or Apple Macintosh version. Release 1.02, Users guide.

Pedersen PB (1992). Engineering analysis concerning energy efficiency standards for refrigerators and freezers in Denmark. Preliminary study. Søborg: dk-Teknik.

Pedersen PH, Lawætz H (1989). LER. *Udvikling, produktion og test af lavenergikøleskabet LER200*. Lyngby: Fysisk Laboratorium III, Danmarks Tekniske Højskole.

Pedersen HB, Pedersen TD, Larsen A (1993). Landbrugets prisforhold 1992/93. Valby: Statens Jordbrugsøkonomiske Institut (Serie C77).

Porskrog H, Olsen O, Sørensen KL, Bjørkman K, Jensen AØ, Møllenbergs, Sørensen S, Pedersen TD (1993). *Økonomien i landbru-*

gets driftrene 1991/92. Valby: Statens Jordbrugsøkonomiske Institut (Serie C76).

Rønning A, Hanssen OJ, Møller H, Gade AL, Haug UC (1993). Livsløpsvurdering av to malingsprodukter. Frederikstad: Stiftelsen Østfoldforskning. Rapportnr. OR.53.93.

SAS Institute (1988). SAS/Stat users guide. Release 6.03 edition. Cary, NC: SAS Institute Inc.

SAS Institute (1990). SAS procedures guide. Cary, NC: SAS Institute Inc.

Tillman A-M, Ekvall T, Baumann H, Rydberg T (1994). Choice of system boundaries in life cycle assessment. L. Cleaner Prod. 2(1):21-29.

van den Hout KD (1994). The impact of uncertainties in impact assessment (abstract). Brussels: SETAC (Abstract Book, Fourth SETAC - Europe Congress "Towards Sustainable Environmental Management", 11-14 April 1994, Brussels, Belgium).

van Hees M (1994). Accuracy of LCA. Paper presented at Fourth SETAC - Europe Congress "Towards Sustainable Environmental Management", 11-14 April 1994, Brussels, Belgium.

Weidema BP (1993). Quality criteria for product life cycle assessment. In Weidema (ed). Environmental assessment of products - A text book on life cycle assessment. Helsinki: UETP-EEE. 2nd edition.

Weidema B P (1994). Qualitative and quantitative parameters in product impact assessment. Pp. 29-35 in Udo de Haes *et al.*: Integrating impact assessment into LCA. Proceedings of the LCA symposium held at the 4th SETAC-Europe Congress, April 11-14, Brussels.

Weidema BP, Pedersen RL, Drivsholm TS (1995). Life cycle screening of food products - Two examples and some methodological proposals. Lyngby: Danish Academy of Technical Sciences.

Winston WL (1996). Simulation modelling using @RISK. Belmont: Duxbury Press.

Young SB, Vanderburg WH (1994). Applying environmental life-cycle analysis to materials. JOM 22-27.

ISSN 0908-9195 ISBN 87-7810-818-7
Pris (inkl. 25% moms): kr. 75,-
Kan købes hos: Miljøbutikken
Telefon: 33 37 92 92 Fax: 33 92 76 90

Miljø- og Energiministeriet **Miljøstyrelsen**
Strandgade 29 · 1401 København K · Telefon 32 66 01 00