

Notat om ændring af metoden for økonomisk benchmarking

Teknisk høring - februar 2024



KONKURRENCE- OG FORBRUGERSTYRELSEN

**Notat om ændring af
metoden for økonomisk benchmarking**

Konkurrence- og Forbrugerstyrelsen
Carl Jacobsens Vej 35
2500 Valby
Tlf.: +45 41 71 50 00
E-mail: kfst@kfst.dk

Ændring af metoden for benchmarking er udarbejdet af
Konkurrence- og Forbrugerstyrelsen.

Februar 2024

Indhold

Kapitel 1	4
Introduktion	4
1.1 Indledning.....	4
1.2 Baggrund og problemstilling	4
Kapitel 2	7
Review af benchmarkingmodeller	7
2.1 Afdækning af forskellige benchmarkingmetoder	7
2.2 Overblik over alternativer til benchmarkingmetoder.....	7
2.3 Evalueringskriterier for et egnet alternativ.....	7
Kapitel 3	10
Beskrivelse af Order-M	10
3.1 Order-M som benchmarkingmodel.....	10

Kapitel 1

Introduktion

1.1 Indledning

Forsyningssekretariatets benchmarking af vandselskaberne har til formål at fastsætte et individuelt effektiviseringskrav for de vandselskaber, der har potentiale for effektivisering. Benchmarkingmetoden justeres løbende med henblik på at opnå en så retvisende benchmarking som muligt. Forsyningssekretariatet har i den forbindelse undersøgt, hvorvidt der kan findes et alternativ til den nuværende benchmarkingmetode, hvor vandselskabernes økonomiske forhold benchmarkes ved brug af en DEA-metode og en SFA-metode.¹

Forsyningssekretariatet har undersøgt alternativer til SFA som benchmarkingmetode, fordi brugen af SFA er forbundet med en række udfordringer i forbindelse med beregningen af den årlige benchmarking af vandselskaberne. Det er i den forbindelse undersøgt, om andre metoder kan håndtere de samme hensyn, som SFA-metoden håndterer i dag. Formålet med undersøgelsen er at afdække, om der findes en mere retvisende, enkel og robust metode til benchmarking.

Forskellige benchmarkingmetoder er forbundet med forskellige fordele og ulemper. Forsyningssekretariatet har derfor vurderet alternativer til en ny benchmarkingmetode ud fra en række evalueringskriterier for, hvad en benchmarkingmodel af vandselskaberne bør leve op til. På grundlag heraf og på grundlag af teoretiske overvejelser og praktiske undersøgelser vurderer Forsyningssekretariatet umiddelbart, at den såkaldte "DEA-Order-M" metode, er den metode, der bedst lever op til de kriterier, der bør stilles til benchmarkingmetoden for vandselskaberne.

DEA-Order-M er en mere enkelt, robust, intuitiv og gennemsigtig udvidelse af DEA-metoden, der altid giver samme eller højere scorer end en standard-DEA. Metoden kombinerer de væsentligste hensyn fra SFA-metoden med hensynene i en standard-DEA-metode. Anvendelse af DEA-Order-M vil derfor kunne erstatte benchmarking med både en DEA og SFA, fordi metoden rummer de samme hensyn, som de to enkeltstående modeller i dag varetager. Der er dermed tale om en forenkling af benchmarkingen i forhold til i dag.

1.2 Baggrund og problemstilling

Forsyningssekretariatet har siden 2011 foretaget benchmarking af de større vand- og spildevandsselskaber med en årlig debiteret vandmængde på over 800.000 m³. I benchmarkingen sammenlignes selskabernes evne til at omsætte omkostninger til produktion. Benchmarkingmodellen udvikles og justeres løbende – dels på grundlag af indberettede oplysninger fra vandselskaberne og dels på grundlag af ny benchmarkingfaglig teori og viden. Benchmarkingen beregner et mål for, hvor effektivt det enkelte selskab er i forhold til de øvrige selskaber.

¹ Yderligere information om benchmarking af vandselskaberne, herunder om DEA-metoden og SFA-metoden, kan findes på Konkurrence- og Forbrugerstyrelsens hjemmeside.

Vandselskaber er naturlige monopoler, eftersom den enkelte forbruger ikke kan vælge et andet vandselskab, end det, som deres ejendom er tilsluttet. Da vandselskaber således ikke agerer på et marked med konkurrence og derfor ikke er udsat for et markedsbestemt effektiviseringspres, foretages benchmarkingen med henblik på at fastsætte individuelle effektiviseringskrav, så vandselskabernes effektiviseringspotentiale indhentes. Benchmarkingen har således til formål at tilskynde vandselskaberne til at have fokus på at være effektive og holde omkostninger nede. Det er til gavn for forbrugere og virksomheder.

I perioden fra 2011 til 2015 under prisloftregulering af vandselskaberne foretog Forsyningssekretariatet årligt den såkaldte *resultatorienterede benchmarking*, hvor vandselskaberne blev benchmarket på driftsomkostninger ved brug af en sammenligningsmetode kaldet Data Envelopment Analysis (DEA).

I 2016 overgik reguleringen af vandselskaberne til økonomiske rammer, og samtidig blev den *totaløkonomiske benchmarking* indført. Det indebærer at vandselskaberne fremover blev benchmarket på både driftsomkostninger og anlægsomkostninger. Siden 2016 har Forsyningssekretariatet foruden en DEA-metode også anvendt en Stochastic Frontier Analysis-metode (SFA). Formålet med at supplere DEA-metoden med en SFA-metode var at kombinere metodernes respektive styrker til at estimere så robuste og retvisende effektiviseringspotentiale som muligt.

Forsyningssekretariatet har løbende opbygget erfaring med at anvende SFA-metoden og har årligt videreudviklet på både DEA- og SFA-metoden. Sekretariatet har også løbende videreudviklet på øvrige elementer i benchmarkingen, som har betydning for de individuelle effektiviseringskrav.

En af styrkerne ved SFA er, at metoden håndterer datastøj. Støj omfatter i denne sammenhæng fx usikkerhed i data, generelle udsving, og hvorvidt frantselskaberne er fuldt repræsentative for resten af sektoren. Denne type af støj tager en standard DEA-metode ikke hånd om, hvilket er årsagen til, at begge modeller har været anvendt.

SFA er mere kompleks end DEA, og det gør SFA mindre gennemsigtig og intuitiv. Komplexiteten i SFA gør blandt andet, at metoden har svært ved at håndtere ændringer i data. Det kan resultere i mindre robuste resultater, hvilket Forsyningssekretariatet vurderer er den største udfordring ved at anvende SFA som benchmarkingmetode for vandselskaberne. Den høje grad af kompleksitet skaber uklarheder og øger risikoen for, at vandselskabers evne til at effektivisere sine omkostninger bliver bedømt enten for højt eller for lavt, når der sammenlignes med en tilsvarende effektivitet i et velfungerende marked. Dertil kommer, at Forsyningssekretariatet i stigende grad har konstateret, at resultaterne fra SFA-metoden er følsomme overfor ændringer i indberettet data. Dette er en konsekvens af, at der indgår relativt få vandselskaber i benchmarkingen, og at 'fronten' i SFA-metoden ikke blot består af de mest effektive vandselskaber, men i stedet for er en vægtning af samtlige vandselskaber i benchmarkingen.

Vandselskaber har løbende givet udtryk for, at SFA-metoden fremstår uigennemsigtig og har en høj grad af kompleksitet. Metoden er særligt kompleks, fordi den både estimerer en omkostningsfunktion ved brug af statistisk regressionsanalyse og indeholder et element, som har til hensigt at identificere fordelingen mellem støj og inefficiens. Når støj og inefficiens skal identificeres, kan SFA-modellen ikke fuldt ud skelne mellem, om ændringer i indberettet data er støj eller ny inefficiens. En følge heraf er fx, at der kan være tilfælde, hvor et vandselskab har reduceret sine omkostninger, men ved anvendelse af SFA-metoden får samme selskab ikke beregnet et tilsvarende lavere effektiviseringspotentiale.

SFA-metoden kræver et datasæt af en vis størrelse for, at der kan beregnes retvisende resultater. SFA-metoden indebærer, at der skal estimeres en lang række centrale parametre, og uden et tilstrækkeligt stort datasæt vil estimererne ikke være robuste eller retvisende. Det er Forsyningssekretariatets vurdering, at med den stigende grad af konsolidering i branchen, vil

resultaterne fra SFA-metoden i de kommende år risikerer at blive estimeret på et datasæt, som indeholder for få observationer til at fastsætte robuste og retvisende effektiviseringspotentialer.

Kapitel 2

Review af benchmarkingmodeller

2.1 Afdækning af forskellige benchmarkingmetoder

Forsyningssekretariatet har undersøgt en række benchmarkingmetoder for at afdække, om metoderne kan anvendes til benchmarking af vandselskaber. Der er foretaget en litteraturgen-
nemgang, og der er foretaget empiriske analyser for så vidt angår nogle af metoderne.

2.2 Overblik over alternativer til benchmarkingmetoder

Forsyningssekretariatet har identificeret otte potentielle benchmarkingmetoder, og sekretari-
atet har undersøgt og evalueret dem nærmere. Alle de undersøgte metoder er fagligt aner-
kendte, teoretisk velfunderede og udbredte inden for benchmarkinglitteraturen. Metoderne
er:

1. Data Envelopment Analysis Order-M (DEA-Order-M)
2. Chance Constrained Data Envelopment Analysis (CCDEA)
3. Fuzzy Data Envelopment Analysis (Fuzzy DEA)
4. Corrected Ordinary Least Squares (COLS)
5. Modified Ordinary Least Squares (MOLS)
6. Semiparametric Stochastic Frontier Analysis (Semiparametrisk SFA)
7. Stochastic Nonlinear Envelopment of Data (StoNED)
8. Key Performance Indicators (KPI)

Fælles for metoderne er, at de bygger videre på eller anvender samme grundlæggende egen-
skaber, som enten DEA-metoden eller SFA-metoden.

For de metoder, der bygger videre på DEA-metoden, omfatter de generelt nogle simple udvi-
delser, der håndterer usikkerhed i data, eller udvidelser, hvor der introduceres et såkaldt sto-
kastisk element til en standard DEA-metode. Fælles for disse metoder er dog, at de er forbun-
det med nogle valg om mængden af datastøj, som ikke er givet på forhånd, og som kan være
vanskelige at fastsætte.

De metoder, der bygger videre på eller anvender samme grundlæggende egenskaber som en
SFA-metoden, er relativt komplekse og derfor svære at gennemskue. Derudover har disse me-
toder til fælles, at de indebærer samme typer af udfordringer, som Forsyningssekretariatet
har erfaret med den SFA-metode, som anvendes i dag. Disse metoder er derfor ikke egnede til
at løse den problemstilling, som har givet anledning til undersøgelserne af mulige alternativer.
Metodernes største svagheder er, at det dels er utydeligt, hvilke incitamenter benchmarkingen
giver til vandselskaberne og dels, at robuste og retvisende resultater er afhængige af, at der er
et tilstrækkeligt stort datasæt.

2.3 Evalueringskriterier for et egnet alternativ

For at vurdere hvilke benchmarkingmetoder, som er mest velegnet i forhold til reguleringen af
vandselskaberne, har Forsyningssekretariatet holdt mulige alternativer op mod fem evalue-
ringskriterier. Evalueringskriterierne omfatter forskellige kvalitative parametre, som Forsy-
ningssekretariatet vurderer er vigtige egenskaber for hensigtsmæssig benchmarking af vand-
selskaberne. De fem evalueringskriterierne er:

1. Faglig robusthed

Faglig robusthed dækker over to elementer; a) præcis og robust estimering af efficiensscorer og b) videnskabelig modenhed.

Ad a)

Præcis og robust estimering af efficiensscorer dækker over de metodiske egenskaber ved en given benchmarkingmetode. Hvis en metode kan estimere efficiensscorer præcist, betyder det blandt andet, at metoden kan tage højde for noget af selskabernes forskellighed. Er metoden robust, betyder det, at efficiensscorerne ikke ændres væsentligt, hvis der sker små ændringer i data.

Ad b)

Videnskabelig modenhed indebærer, hvor udbredt en metode er i de videnskabelige miljøer. En af de parametre, som kan bestyrke en metodes faglige robusthed, er, at den er udbredt i videnskabelige kredse og har været anvendt i det videnskabelige miljø over en periode, samt at den er anvendt af andre regulatorer.

2. Tilgængelighed

Metoden skal bygge på teori og praksis, som er forståelig og anvendelig. Der er store forskelle på, hvor komplekse forskellige metoderne er, og derfor hvor mange ressourcer og hvor meget ekspertviden der kræves for at forstå og arbejde med metoderne. Af hensyn til kommunikation med branchen, fremdrift i benchmarkingarbejdet, herunder i forbindelse med løbende tilpasninger og justeringer, er det en fordel, at en metode i udgangspunktet er mere enkelt.

3. Tydelige incitamenter

En benchmarkingmetode skal give selskaberne et incitament til at være omkostningseffektive. Idet formålet med benchmarkingen er at give selskaberne incitament til en omkostningsminimerende adfærd, er det en fordel, at modellen er så intuitiv som muligt. Er modellen intuitiv, er det lettere at kommunikere om metoden og metodens resultater og derved tydeliggøre incitamenterne. Eksempelvis når et selskab for en given produktion reducerer sine omkostninger med 100.000 kr., bør en intuitiv benchmarking også afspejle, at selskabet har et potentiale til at effektivisere sig, der er tilsvarende 100.000 kr. mindre end før.

4. Metodens praktiske egenskaber

En velegnet benchmarkingmetode skal som udgangspunkt kunne anvende nuværende tilgængelige data. Det skyldes, at selskabernes indberetning af ny data både kræver ressourcer hos selskaberne og hos Forsyningssekretariatet. Det er derfor en fordel, hvis en model kan fungere med eksisterende tilgængelige data eller med simple udvidelser eller justeringer. Nogle metoder kræver også, at en computer har regnekraft i et omfang, som ikke er tilgængelig på normale computere. Derfor kan selve beregningen af efficiensscoren være en udfordring for nogle benchmarkingmetoder.

5. Fremadrettede anvendelsesmuligheder

Det skal være muligt fremover at implementere flere og nye variable samt yderligere rammebetingelser i benchmarkingmetoden. Det skyldes, at der løbende kommer adgang til nye variable. Derudover kan det være hensigtsmæssigt, at en ny benchmarkingmodel kan håndtere ændringer eller tilføjelser til lovgivningen på området. Derfor bør en metode have fleksibilitet til, at den kan udvikles til at imødekomme fremtidige behov.

Vurderingen af de otte benchmarkingmetoder fremgår neden for af tabel 2.1, hvor metoderne har fået farvekode grøn, gul eller rød, alt efter i hvilken grad metoderne vurderes at leve op til evalueringskriterierne.

En mere detaljeret gennemgang af hver metode og en vurdering af, i hvilket omfang hvert alternativ opfylder de fem evalueringskriterier, kan læses i "Bilag 1 – Review af benchmarkingmetoder".

Tabel 2.2 Evaluering af alternativet til benchmarkingmetode

Evalueringsskriterier:	Faglig robusthed	Tilgængelighed	Tydelige incitamenter	Metodens praktiske egenskaber	Fremadrettede anvendelsesmuligheder
Order-M					
CCDEA					
Fuzzy DEA					
COLS					
MOLS					
Semiparametrisk SFA					
StoNED					
KPI					

Anm.: Tabellen viser, i hvilken grad benchmarkingmetoderne vurderes at leve op til evalueringsskriterierne. Grøn indikerer, at metoden i høj grad lever op til evalueringsskriterieret. Gul indikerer, at der er usikkerhed, men at det vurderes, at det kan håndteres. Rød indikerer, at der er væsentlige udfordringer med metoden, og at den derfor i ringe grad eller slet ikke lever op til evalueringsskriterieret.

Kilde: Egen tilvirkning

På baggrund af undersøgelser af metoderne og vurdering af metoderne i forhold til evalueringsskriterierne vurderes det, at DEA-Order-M (herefter forkortet *Order-M*) er den mest velegnede benchmarkingmetode til brug for regulering af vandselskaberne. Metoden er en af de mindst komplekse metoder, og den er relativ intuitiv og tilgængelig. Order-M er en simpel udvidelse af standard DEA-metoden. DEA er i forvejen relativt intuitiv at forstå, og en udvidelse heraf, som også er af relativt lav kompleksitet, indebærer tydelige incitamenter til vandselskaberne. Udvidelsen gør desuden, at DEA-metodens egenskaber tilføjes en dimension, som indebærer, at der tages højde for datastøj, hvilket er en af de primære styrker ved SFA-metoden.

I tillæg er resultaterne fra Order-M i mindre grad afhængig af oplysninger fra enkelte selskaber. Det giver metoden gode praktiske egenskaber i forhold til at sikre fremdrift i benchmarkingen, selvom der skulle ske løbende ændringer eller opdateringer af selskabernes data. Desuden er Order-M fleksibel i forhold til mulighederne for fremover at kunne inkludere nye variable i benchmarkingen.

Kapitel 3

Beskrivelse af Order-M

3.1 Order-M som benchmarkingmodel

DEA-Order-M (Order-M) er en simpel udvidelse af den standard DEA-metode, som Forsyningssekretariatet anvender i dag. Order-M udvider DEA ved at udjævne fronten. Udjævnningen af fronten sker ved, at Order-M lader flere selskaber indgå i frontfastsættelsen i forhold til DEA. I Forsyningssekretariatets benchmarking af vandselskaberne har fronten i DEA typisk bestået af 3-5 efficiente selskaber (fuldt effektive). Det betyder, at retvisende resultater ved anvendelse af DEA i høj grad afhænger af, at disse få efficiente selskaber indberetter korrekte data, og at de er repræsentative for hele sektoren. I Order-M består fronten imidlertid typisk af mere end de 20 mest effektive selskaber. Order-M er derfor mere robust overfor udsving i data og for spørgsmålet om, hvorvidt enkelte selskaber er repræsentative for hele sektoren.

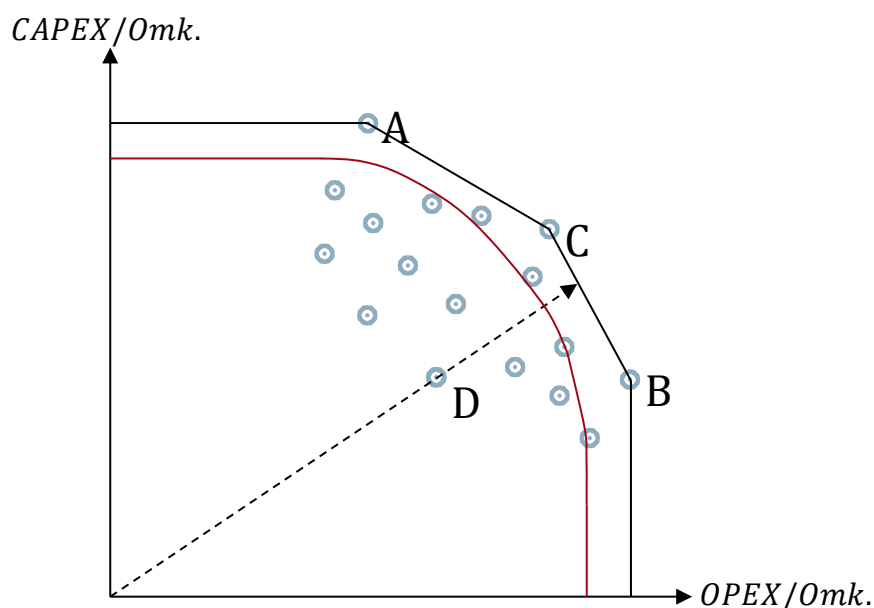
Order-M beregnes ved at lave en DEA-model over flere omgange (herefter kaldet iterationer). I hver iteration trækkes M-antal tilfældige selskaber, som kan være med til at danne fronten i den givne iteration. Dette er i modsætning til en standard DEA, hvor alle selskaber som udgangspunkt kan være med til at danne fronten. Da der i hver iteration udtages M-tilfældige selskaber, er der sandsynlighed for, at de efficiente selskaber ikke udtages og derfor heller ikke danner fronten i den givne iteration. Hvis de efficiente selskaber ikke er med i udtrækket, vil de næstmest effektive selskaber være med til at definere fronten i den givne iteration. Den endelige front, som selskaberne sammenlignes med i Order-M, er et gennemsnit af resultaterne fra alle iterationerne.

Order-M stiller altid selskaberne bedre end DEA, jf. Figur 3.1. Figuren sammenligner en Order-M-metode med en standard DEA-metode. De sorte linjestykker angiver fronten i DEA, og de røde linjer angiver fronten i Order-M. De sorte punkter angiver alle selskaberne.

Figuren viser, at DEA-fronten i dette eksempel består af de tre efficiente selskaber, og at alle andre selskaber ligger "under" fronten. Det vil sige, at alle andre selskaber er ineffektive, hvilket måles som afstanden fra det enkelte selskab til fronten. Fronten i Order-M ligger altid under DEA-fronten og består som tidligere beskrevet af et gennemsnit af flere iterationer. I de mest ekstreme iterationer vil fronten være identisk med DEA-fronten, imens den i andre iterationer vil ligge lavere. Afstanden fra de enkelte selskaber til fronten er dermed kortere end til fronten i en standard DEA-metode. Selskaber, der ligger nordøst for fronten i Order-M, vil have en negativ afstand til fronten og blive anset som såkaldt superefficiente.²

² Superefficiente selskaber har en efficiensscore, der er større end 1. Det betyder, at de klarer sig bedre end modellen finder muligt. Superefficiens er derfor et resultat af støj i modellen.

Figur 3.1 Sammenligning af fronten i DEA og Order-M



Anm.: Figuren sammenligner fronten i en DEA og i en Order-M. Hvert punkt indikerer et selskab. De sorte optrukne linjer angiver DEA-fronten, og de røde linjer angiver Order-M-fronten. Figuren er lavet på simuleret data.

Kilde: Egen fremstilling

Frontudjævningen i Order-M er udtryk for, at modellen tager hensyn til støj. Netop hensynet til støj har været en vigtig argumentation for, at Forsyningssekretariatet tidligere har anvendt både DEA og SFA, da SFA tager hensyn til støj, hvilket DEA ikke gør. Da Order-M tager hensyn til støj, kan den erstatte både DEA og SFA, så Forsyningssekretariatet fremadrettet kun bruger én model.

Order-M bygger på alle de samme antagelser som DEA-metoden. Det betyder, at der ikke er behov for ændringer i det data, som selskaberne indberetter, selvom der overgås til denne nye benchmarkingmetode.

For at implementere Order-M er det dog nødvendigt at fastsætte størrelsen af M – det vil sige, hvor mange selskaber, der udtrækkes tilfældigt i hver iteration.

Når M er stor, vil der være stor sandsynlighed for, at de efficiente selskaber danner front i hver iteration. Det betyder, at Order-M-fronten vil ligne en standard DEA, når M er høj. Når M omvendt er lav, er det mindre sandsynligt, at de efficiente selskaber danner front i hver iteration, og derfor er det også mindre sandsynligt, at Order-M-fronten ligner en standard DEA. På den måde har M betydning for, hvordan fronten i Order-M ser ud.

Fastsættelsen af det konkrete M kræver i sig selv forskellige overvejelser. Det er en følge af, at fastsættelsen af M er udtryk for størrelsen på mængden af støj, der skal indarbejdes i metoden. Med henblik på at muliggøre fastsættelsen af et konkret M har Forsyningssekretariatet opstillet en række mulige egenskaber, som Order-M i størst muligt omfang bør opfylde.

Disse egenskaber er i tillæg til de evalueringskriterier, som Order-M også opfylder for at være et egnet alternativ til SFA-metoden. Egenskaberne er som følgende:

1. Resultater er ikke følsomme over for data fra enkelte selskaber
2. Resultater tager hensyn til støj uden at udvande formålet om identificering af efficiens
3. Fronten udgøres af mange selskaber, så fronten har høj grad af fleksibilitet
4. Modellen opfylder betingelser og antagelser fra DEA

Egenskaberne og fastsættelsen af M diskuteres i "Bilag 2 - Teknisk gennemgang af Order-M" sammen med en mere teknisk gennemgang af Order-M.

Review af metoder for benchmarking

Teknisk bilag 1

Februar 2024



KONKURRENCE- OG FORBRUGERSTYRELSEN

Review af metoder for benchmarking

Konkurrence- og Forbrugerstyrelsen

Carl Jacobsens Vej 35
2500 Valby
Tlf.: +45 41 71 50 00
E-mail: kfst@kfst.dk

Review af benchmarkingmetoder er udarbejdet af
Konkurrence- og Forbrugerstyrelsen.

Februar 2024

Indhold

Kapitel 1	4
Indledning	4
1.1 Baggrund	4
Kapitel 2	5
DEA-Order-M (Order-M)	5
Kapitel 3	7
Chance Constrained Data Envelopment Analysis (CCDEA)	7
Kapitel 4	10
Fuzzy DEA	10
Kapitel 5	13
Corrected Ordinary Least Squares (COLS)	13
Kapitel 6	16
Modified Ordinary Least Squares (MOLS)	16
Kapitel 7	19
Semiparametric Stochastic Frontier Analysis (Semiparametrisk SFA)	19
Kapitel 8	21
Stochastic Nonlinear Envelopment of Data (StoNED)	21
Kapitel 9	23
Key Performance Indicators (KPI)	23
Kapitel 10	25
Referencer	25

Kapitel 1

Indledning

1.1 Baggrund

Forsyningssekretariatet foretager i dag en årlig benchmarking af vandselskaberne, hvor selskabernes økonomiske forhold sammenlignes ved brug af hhv. en DEA-metode og en SFA-metode. Begge metoder er fagligt anerkendte og bredt anvendt både i akademiske kredse og af diverse regulatorer af forsyningsselskaber.

I modsætning til DEA er SFA baseret på mere komplekse modelantagelser, som besværliggør gennemsigtheden og den intuitive forståelse af metoden bag SFA-metoden. Komplexiteten i SFA gør blandt andet, at metoden har svært ved at håndtere ændringer i data. Herudover har vandselskaberne løbende givet udtryk for, at SFA-metoden har en høj grad af kompleksitet og fremstår uigennemskuelig. En anden udfordring med SFA-metoden er, at metoden stiller krav til at have et tilstrækkeligt stort og ensartet datasæt for at kunne beregne robuste og retvisende effektiviseringskrav.

Derfor har Forsyningssekretariatet undersøgt en række af mulige alternativer til det nuværende set-up for benchmarkingen, hvor der både anvendes DEA og SFA. De alternative benchmarkingmetoder er vurderet ud fra en række evalueringskriterier:

- » Faglig robusthed
- » Tilgængelighed
- » Tydelige incitamenter
- » Metodens praktiske egenskaber
- » Fremadrettede anvendelsesmuligheder

Ovenstående evalueringskriterier er nærmere beskrevet i *Forslag til ændring af metoden for benchmarking*.

En del af de alternative metoder tager udgangspunkt i metoderne bag DEA- og SFA, der i dag anvendes til fastsættelse af det individuelle effektiviseringskrav i reguleringen. Baggrundsnoter om metoderne bag Forsyningssekretariatets nuværende DEA- og SFA-model kan læses på Forsyningssekretariatets hjemmeside [her](#).

Kapitel 2

DEA-Order-M (Order-M)

DEA-Order-M er en simpel udvidelse af DEA. Order-M udvider DEA ved at udjævne fronten. Udjævnningen af fronten sker ved, at Order-M lader flere selskaber indgå i frontfastsættelsen i forhold til en standard DEA-metode. Frontudjævnningen i Order-M kan betragtes som, at modellen tager hensyn til støj. I Forsyningssekretariatets benchmarking af vandselskaberne består fronten i DEA typisk af de 3-5 efficiente selskaber.¹ I Order-M består fronten imidlertid typisk af mere end de 20 mest effektive selskaber. Order-M er derfor mere robust overfor udsving i data og for, om enkelte selskaber er repræsentative for hele sektoren.

Order-M blev første gang foreslået som metode til at beregne efficiensscorer i 2002² og er siden blevet videreudviklet inden for litteraturen.³

Order-M beregnes ved at lave en DEA-model over flere iterationer. I hver iteration trækkes kun en andel (M antal) af det samlede antal selskaber, som kan være med til at danne fronten i den givne iteration. På den måde kan også de mindre effektive selskaber⁴ være med til at danne fronten i Order-M, hvis de efficiente selskaber ikke er udtrukket i den enkelte iteration. Dette er i modsætning til en standard DEA, hvor alle selskaber indgår, og det dermed kun er de efficiente selskaber, der danner fronten. Da der i hver iteration udtages M tilfældige selskaber, er der sandsynlighed for, at de efficiente selskaber ikke udtages, og derfor heller ikke danner fronten i den givne iteration. Den endelige front, som selskaberne sammenlignes med i Order-M, er et gennemsnit af resultaterne fra alle iterationerne.

For at implementere Order-M er det dog nødvendigt at fastsætte størrelsen af M – det vil sige, hvor mange selskaber der udtrækkes tilfældigt i hver iteration. Størrelsen af M har betydning for sandsynligheden for at trække bestemte selskaber i hver iteration. M har derfor betydning for, hvordan fronten i Order-M ser ud. Det er dog ikke entydigt at fastsætte et konkret M, fordi det kræver et valg om mængden af støj, der skal indarbejdes i metoden. I den videnskabelige litteratur er der forskellige tilgange, men for alle tilgange skal der i sidste ende træffes et valg. Selvom flere af tilgangene er statistisk funderet, vil der altid i sidste ende træffes en beslutning, der må bero på de statistiske mellemresultater.⁵

¹ 'Efficiente selskaber' betegner de vandselskaber som har været bedst til at omsætte omkostninger til produktion eller service, og er de (front)selskaber som de øvrige selskaber derfor måles op imod i en standard DEA-model.

² Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, s. 1-25.

³ Se fx Da Silva, T., Martins-Filho, C., & Ribeiro, E. (2016). Comparison of nonparametric efficiency estimators: DEA, FDC, DEAC, FDCH, order-m and quantile. *Economic Bulletin*, 36(1), s. 118-131.

⁴ 'Mindre effektive selskaber' betegner de vandselskaber som ikke har formået – i lige så høj grad som frontelskaberne – at omsætte omkostninger til produktion eller service.

⁵ Daraio & Simar (2007) viser, at DEA-order-M har de samme statistiske egenskaber som DEA. Herudover viser de, at DEA-order-M er mindre påvirkelig for antallet af variable og generelt en mere robust estimator for efficiensscorer.

**Boks 2.1
Order-M****Evaluering af metoden****Faglig robusthed**

Order-M er afhængig af det valgte M og kan ikke modellere støj på samme måde som ved de statistiske metoder som eksempelvis SFA.

Order M indebærer dog ikke, at der skal foretages et arbitrært valg om fordelingen af støj i data. Metoden anvender det konkrete data til at udjævne DEA-fronten og på den måde indføres et element, som håndterer fordelingen af støj ud fra de observerede værdier.

Det er dog ikke entydigt hvordan parameteren M fastsættes. Selvom der er forskellige forslag i litteraturen til at fastsætte denne, er der endnu ikke konsensus om én metode. Forsyningssekretariatet har på baggrund af empiriske analyser dog fundet frem til, at resultaterne fra Order M er relativt robuste i forhold til fastsættelsen af et specifikt M. Metoden er derfor stadig relativ robust, da parameteren M kan fastsættes i et bredt spænd uden at det medfører usikkerhed i resultaterne.

Forsyningssekretariatet har ikke kendskab til, at Order-M anvendes af andre regulatorer. I dag bliver Order-M hovedsageligt brugt i forbindelse med at finde outliers eller vurdere, hvor stor indflydelse enkelte frontskaber har for en model.

Tilgængelighed

Order-M er en simpel udvidelse af DEA. Det betyder, at metoden relativt nemt kan implementeres og anvendes som ny benchmarkingmetode.

Tydelige incitamenter

Order-M giver et stærkt incitament til at være omkostningseffektiv. Selskabernes aktiviteter slår igennem 1:1 i modellen,⁶ og det er derfor nemt for selskaberne at se metodens incitament, når reduktion i omkostninger også tilsvarende reducerer selskabet effektiviseringspotentiale.

Metodens praktiske egenskaber

Der er ingen praktiske udfordringer med at implementere metoden. Metoden kan anvendes med de data og software, som allerede anvendes i benchmarkingen i dag.

Fremadrettede anvendelsesmuligheder

Order-M kan relativt nemt tilpasses ved modelændringer, eksempelvis kan metoden håndtere flere inputs og outputs, end det Forsyningssekretariatet har anvendt hidtil i benchmarkingen.

⁶ Dette gælder kun for selskaber, der ikke er efficiente. Efficiente selskaber får ikke et bedre resultat af at blive mere effektive, så længe de identificeres som efficiente.

Kapitel 3

Chance Constrained Data Envelopment Analysis (CCDEA)

Chance Constrained Data Envelopment Analysis (CCDEA) er en udvidelse af en traditionel DEA. CCDEA tager, i modsætning til en standard DEA, højde for, at der kan være usikkerhed i frontfastsættelsen. Det gøres gennem en stokastisk forskydning af fronten, som delvist korrigerer for datausikkerhed. Ud over forskydningen af fronten er CCDEA og DEA metodisk identisk.

CCDEA blev forslået første gang i 1993⁷ og er siden løbende blevet udviklet i flere forskellige retninger, hvori der gøres forskellige antagelser.⁸ Der er ikke videnskabelig konsensus om, at én af CCDEA-modellerne er den optimale.

CCDEA opdeler variable i to kategorier: stokastiske og deterministiske variable.⁹ Stokastiske variable dækker over variable, som antages at indeholde usikkerhed, mens deterministiske variable ikke indeholder usikkerhed. Ofte vælges enten alle inputs eller alle outputs at være stokastiske, men der kan også vælges en kombination.

Frontfastsættelsen i CCDEA sker herefter efter næsten samme metode som en almindelig DEA. Forskellen mellem de to opstår først, idet der sker en forskydning af CCDEA-fronten på baggrund af den estimerede støj i det underliggende data, jf. Figur 2.

Figur 2 viser sammenhængen mellem fronten i DEA og CCDEA med ét deterministisk input og ét stokastisk output. Selskab A er ineffektivt og bliver benchmarket mod punktet A' i en standard DEA model. I CCDEA antages det, at outputmængderne i punkterne A og A' er behæftet med usikkerhed. Den observerede outputmængde i A er dermed ikke nødvendigvis den mængde, selskab A forventes at kunne producere, hvis de skulle gentage deres produktion med samme inputs. Figuren viser også, at usikkerheden omkring de producerede outputmængder i A og A' er normalfordelt (de røde markeringer).¹⁰ Det betyder, at outputmængderne, ved en gentagelse af produktionsprocessen, kan være et andet sted inden for denne fordeling.

Sandsynligheden for at outputmængden ændrer sig ved gentagen produktion, danner grundlag for CCDEA-frontens forskydning. Forskydningen foretages, så den forventede outputmængde for selskab A med en given sandsynlighed ligger under det forventede punkt for A'.

⁷ Land, K. C., Lovell, C. A., & Thore, S. (1993). Chance-Constrained Data Envelopment Analysis. *Managerial and Decision Economics*, 14(6), s. 541-554.

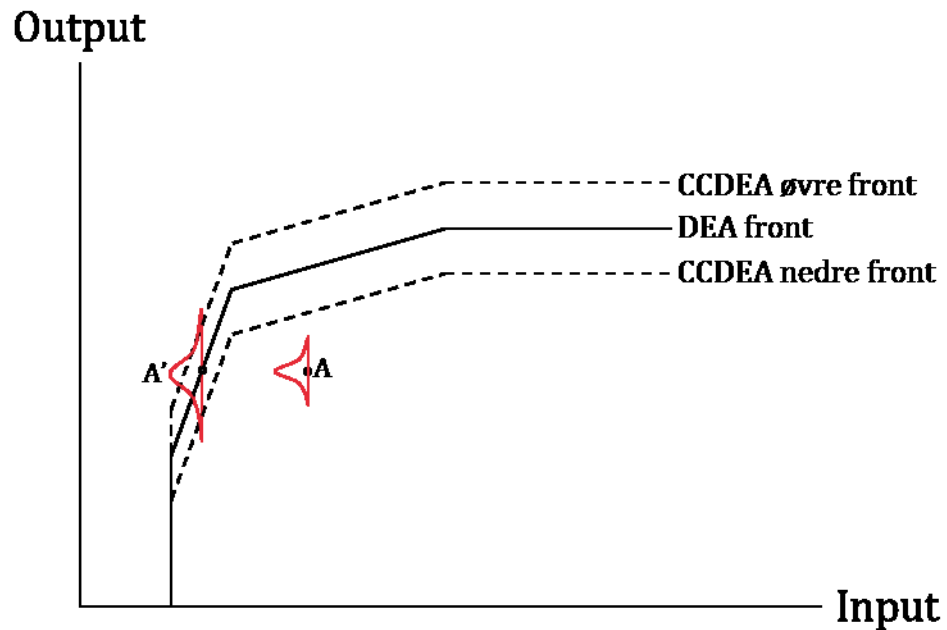
⁸ Se fx Dyson, R. G., & Estelle, A. S. (2010). Data envelopment analysis, operational research and uncertainty. *Journal of the Operational Research Society*, 61(1), s. 25-34. og Olesen, O. B. (2006). Comparing and combining two approaches for chance constrained DEA. *Journal of Productivity Analysis*, 26(2), s. 103-119.

⁹ En almindelig DEA antager, at alle variable er deterministiske (Bogetoft & Otto, *Benchmarking with DEA, SFA, and R*, 2011).

¹⁰ Her antages at der skal bruges en normalfordeling og at det observerede punkt A ligger i midten af fordelingen. Det er ikke nødvendigvis tilfældet i virkeligheden.

Hvis man ønsker at være forholdsvis sikre på, at selskab A ikke benchmarkes mod et urealistisk højt niveau, kan sandsynligheden sættes til 95 %. Det betyder, at der er 95 % sandsynlighed for, at forskellen mellem selskab A og A', som er efficient, reelt skyldes inefficiens og ikke blot tilfældigheder. Hvis sandsynligheden sættes til at være høj, vil fronten med stor sandsynlighed blive forskudt nedad i forhold til DEA fronten. Tilsvarende kan fronten forskydes opad, hvis sandsynligheden sættes lavt, hvilket dog sjældent gøres.

Figur 2 Chance-Constrained DEA



Anm.: Fronten i CCDEA forskydes enten op eller ned afhængig af datausikkerheden og den tilladte afvigelse i de stokastiske variable.

Kilde: Egen tilvirkning

For at CCDEA kan forskyde fronten, kræver det, at fordelingen for alle observationer estimeres (tilsvarende den røde normalfordeling for selskab A i Figur 2). Det kræver tre primære antagelser:

1. *Det underliggende data for et selskabs output (og/eller input) antages at være normalfordelt.* I praksis bør det undersøges, om andre fordelinger er mere passende. Mange studier antager, at alle selskaber har samme varians på deres underliggende output og/eller input.
2. *Den observerede værdi antages at være den mest sandsynlige værdi.* Det betyder fx, at et selskabs observerede outputmængde antages at være den mest sandsynlige outputmængde for selskabet. Reelt er det dog muligt, at selskabets observerede outputmængde, af forskellige årsager, er unormalt høj eller lav.
3. *Afhængigheden mellem selskabernes stokastiske variable skal bestemmes.* Da fordelingen af de stokastiske variable sammenlignes med hinanden kræver det, at afhængigheden (kovariansen) mellem selskabernes produktion er kendt. Hvis selskabernes outputmængder afhænger af hinanden, påvirker det sammenligningen.

**Boks 3.1
CCDEA****Evaluering af metoden****Faglig robusthed**

CCDEA er grundlæggende ligeså fagligt robust som en standard DEA, da de to næsten er identiske, for så vidt angår modellernes generelle egenskaber.

Til brug for frontforskydningen i CCDEA kræves dog en række centrale antagelser, jf. punkt 1-3 ovenfor, for at danne en effektiv front. Herunder antagelse om sandsynligheden for, at det observerede data er det mest sandsynlige. Validiteten af disse antagelser kan ofte ikke bekræftes empirisk, idet der ikke er metoder hertil. Derfor er metoden behæftet med stor usikkerhed i forhold til beregning af efficiensscorer.

Metoden blev udviklet i 1990'erne, og har siden været anvendt i den videnskabelige litteratur, og er en af de mest anvendte Stokastiske DEA-metoder. Forsyningssekretariatet har dog ikke kendskab til, at den anvendes af andre regulatorer.

Tilgængelighed

CCDEA er en simpel udvidelse af DEA. Det betyder, at metoden relativt nemt kan implementeres og anvendes som ny benchmarkingmetode. Det vil dog kræve flere ressourcer at analysere hvilken CCDEA-model og hvilke antagelser, der passer til vandsektoren.

Tydelige incitamenter

CCDEA giver ikke lige så stærke incitamenter til at være omkostningseffektiv, fordi selskabernes aktiviteter ikke slår igennem 1:1, eksempelvis når et selskab reducerer deres omkostninger.

Metodens praktiske egenskaber

Metoden kan anvendes med de data og software som allerede anvendes i benchmarkingen i dag.

Fremadrettede anvendelsesmuligheder

Generelt gælder det, at metoden har samme fremadrettede anvendelsesmuligheder som en DEA-model. CCDEA er derfor også begrænset af antallet af frihedsgrader i samme omfang som en standard DEA-model. Det betyder, at flere variable som udgangspunkt medfører højere efficiensscorer.

Kapitel 4

Fuzzy DEA

Fuzzy DEA anvender såkaldt *fuzzy logik* til at tage højde for usikkerhed i data og modelantagelserne i en standard DEA. Fuzzy logik anskuer data i intervaller frem for konkrete værdier. Det betyder, at der for hvert datapunkt ikke findes en enkelt "sand" værdi, men derimod en række værdier i et interval, som alle kan være "sande". Der er mange varianter af fuzzy DEA, men Forsyningssekretariatet har undersøgt metoden bag den mest udbredte, den såkaldte *alpha-cut* metode.

Fuzzy logik blev første gang forslået til brug i DEA i Sengupta (1992 A) og Sengupta (1992 B). Metoden har været bredt anvendt i litteraturen sidenhen og har udviklet sig i flere forskellige retninger.¹¹

For at anvende fuzzy DEA skal der på forhånd fastsættes intervaller omkring det observerede data, hvori den "sande" værdi ligger. Der er ikke én konkret metode til at fastsætte intervaller, hvilket har resulteret i, at de fleste studier bruger ekspertudsagn til at estimere intervallerne. I vandsektoren kan samarbejde med branchen eller et udvalg af eksperter som har indsigt i vandsektoren, bruges til at vurdere usikkerheden i selskabernes indberettede data. Ud fra vurdering kan der laves et interval, der er stort nok til, at den sande værdi med sikkerhed ligger i intervallet. Samtidig skal man dog være opmærksom på, at intervallet ikke bliver unødvendigt stort, fordi det vil udvande målet om efficiens.

Når intervallerne er estimeret kan de implementeres i et DEA-setup, jf. Figur 4.1. Det observerede data er givet ved punkt A, B, C og D. De horisontale røde linjer angiver intervallet for selskabernes input og de vertikale røde linjer angiver intervallet for selskabernes output. De røde kasser angiver dermed alle selskabernes mulige kombinationer af inputs og outputs. De sorte optrukne linjestykker angiver en standard DEA-front. For at fastsætte efficiensscore, anvendes en såkaldt Alpha-parameter, som skal sættes til en værdi mellem 0 og 1. Når Alpha er lig 1 er fuzzy DEA og en standard DEA identisk, men i takt med at Alpha-værdien sænkes til 0, dannes et bånd mellem to yderpunkter, det ene punkt er den såkaldte "umulige efficiensscore" og det andet punkt den såkaldte "risikofrie efficiensscore".

Den umulige efficiensscore

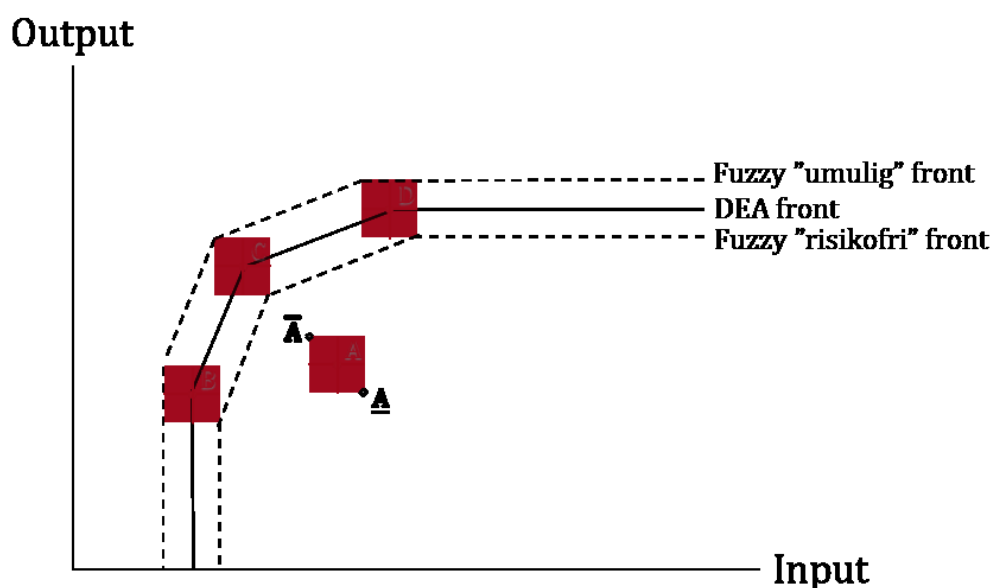
Den såkaldte umulige efficiensscore beregner den laveste mulige efficiensscore for selskabet givet værdien af alpha. Hvis alpha er 0, bruges det højeste input og laveste output i selskab A's interval, hvilket er de værdier, som stiller selskabet dårligst, jf. punkt A i Figur 4.1. Samtidig bruges de værdier, som stiller de resterende selskaber bedst, og dermed skaber den mest restriktive front. Det vil sige det laveste input og højeste output, jf. den øvre stiplede linje "Fuzzy umulig" front" i figuren.

¹¹ Se fx Arana-Jiménez, M., Sánchez-Gil, M. C., & Lozano, S. (2020). Efficiency Assessment and Target Setting Using a Fully Fuzzy DEA Approach. *International Journal of Fuzzy Systems*, 4, s. 1056-1072. og Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: two decades in the making. *European journal of operational research*, 3, s. 457-472.

Den risikofrie efficiensscore

Den risikofrie efficiensscore er den modsatte af den umulige efficiensscore. Det vil sige, at man stiller selskabet bedst muligt givet værdien af α . Når α er lig 0 bruger man det højeste output og laveste input for selskab A ($\bar{A}$), og laveste output og højeste input for de resterende selskaber (den nedre striplede linje).

Figur 4.1 Fuzzy DEA



Anm.: Fuzzy DEA bruger intervaller omkring det observerede data at tage højde for usikkerhed. $\bar{A}$ sammenlignes med Fuzzy "umulig" front og $\underline{A}$ sammenlignes med Fuzzy "risikofri" front. De tilhørende efficiensscore betegnes som de mest ekstreme i hver sin retning

Kilde: Egen tilvirkning

Fuzzy DEA giver dermed et interval af efficiensscore fra den "umulige" til den "risikofrie". Der findes forskellige bud på at vægte disse efficiensscore sammen til en samlet efficiensscore, men metoden er overordnet kun tænkt til at give et interval af efficiensscore.

Boks 4.1
Fuzzy DEA**Evaluering af metoden****Faglig robusthed**

Fuzzy DEA kræver, at der på forhånd er kendskab til usikkerheden (intervallet) omkring det observerede data. Det har stor betydning for modellens resultater, men er vanskeligt at estimere. Herudover resulterer metoden i flere efficiensscorer, og det er ikke entydigt, hvordan resultaterne skal omsættes til effektiviseringskrav i regulering af vandsektoren.

Fuzzy DEA er fortsat under udvikling i den videnskabelige litteratur, hvilket blandt andet betyder, at der er mange forskellige varianter og ingen konsensus om, at én variant er bedre end de andre. Forsyningssekretariatet har ikke kendskab til, at metoden anvendes af andre regulatorer

Tilgængelighed

Den version af fuzzy DEA der er præsenteret i dette bilag er den mest simple fuzzy DEA-metode. Denne version er let at implementere, men andre versioner er mere vanskelige, og kan være komplekse. I tillæg bliver metoden teoretisk vanskelig, når der skal anvendes statistiske metoder, til at fastsætte intervallet rundt om det observerede data samt kombinere efficiensscorerne til et enkelt mål.

Tydelige incitamenter

Modellens incitamenter afhænger af den variation, der anvendes, samt hvordan intervallet omkring den observerede værdi fastsættes. Det er muligt at opsætte metoden, så der er de samme tydelige incitamenter som i en standard DEA. Der er dog også risiko for, at metoden forveksler reelle effektiviseringer med støj, og det medfører, at selskaberne ikke fremstår lige så effektive som de burde.

Metodens praktiske egenskaber

Metoden kan som udgangspunkt bruges med de data og software som allerede anvendes i benchmarkingen i dag. Der kan dog opstå et behov for, at der skal indsamles yderligere data, hvis de intervaller, der skal bruges i modellen, vurderes at skulle være selskabsspecifikke i den konkrete implementering af metoden.

Fremadrettede anvendelsesmuligheder

Generelt gælder, at metoden har samme fremadrettede anvendelsesmuligheder som en DEA-model. Fuzzy DEA er begrænset af antallet af frihedsgrader i samme omfang som en DEA-model. Det betyder, at flere variable som udgangspunkt fører til højere efficiensscorer.

Kapitel 5

Corrected Ordinary Least Squares (COLS)

Corrected Ordinary Least Squares (COLS) er en statistisk funderet benchmarkingmetode, hvor der med OLS estimeres en omkostningsfunktion, som ”korrigeres” (forskydes) således, at ingen selskaber ligger foran omkostningsfunktionen. Herefter måles alle selskaber op imod omkostningsfunktionen. En omkostningsfunktion estimerer de minimale omkostninger, som en virksomhed kan opretholde en given produktion ved, hvor en OLS-estimeret omkostningsfunktion antager, at alle selskaber bruger de minimale omkostninger. COLS er første gang præsenteret i 1957¹² og har været anvendt til økonomisk benchmarking siden.

COLS estimerer efficiens i to trin:

1. Først estimeres en almindelig OLS-model, hvor der beregnes en front ved brug af en regressionsanalyse.
2. Herefter parallelforskydes fronten nedad, indtil alle residualer er mindst 0 – det vil sige, at fronten forskydes så alle selskaber enten er bagved fronten eller ligger på fronten med fronten. Fronten danner nu benchmarket for alle selskaber, og kun de selskaber som er ligge på fronten vil være fuldt ud efficiente. Det betyder, at COLS er deterministisk i sin estimering af efficiens, selvom den er parametrisk i sit udgangspunkt.

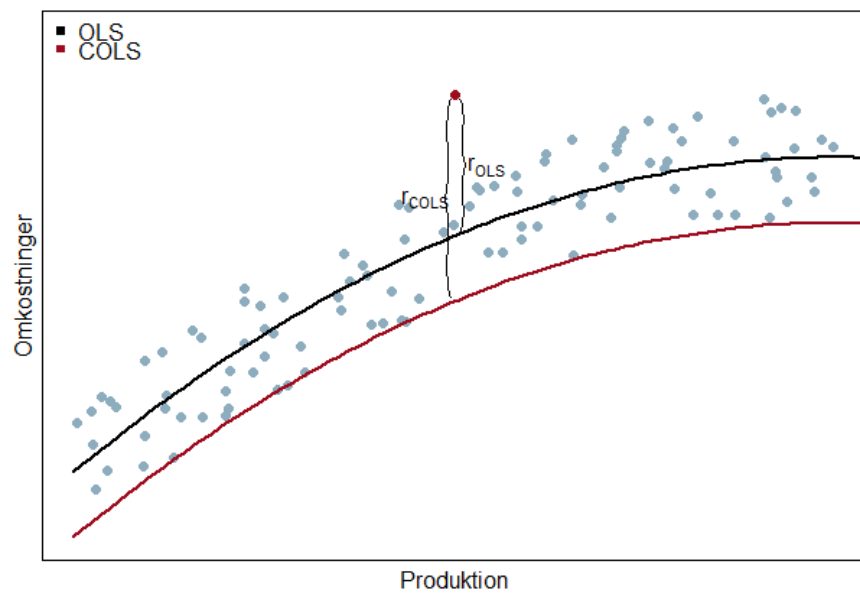
COLS og OLS er illustreret i Figur 5.1. Figuren viser to omkostningsfunktioner for hhv. OLS (den sorte linje) og COLS (den røde linje). Det ses, at COLS alene er en parallelforskydning af OLS-omkostningsfunktionen nedad, indtil alle datapunkter ligger over omkostningsfunktionen, dvs. alle residualer, r_{COLS} , er positive eller 0. r_{COLS} er dermed effektiviseringspotentiallet for selskaberne.

Idet COLS er en mere enkel forskydning af omkostningsfunktionen, kan der opstå en situation, hvor omkostningsfunktionen krydser x-aksen. Det betyder, at selskaber, som benchmarkes med den negative del af omkostningsfunktionen, vil have et effektiviseringspotentialt på over 100 pct., idet de skal have negative omkostninger for at være efficiente.

Metoden er også følsom overfor store udsving i selskabsstørrelse, fordi COLS er en mere enkel forskydning af omkostningsfunktionen og hele afstanden mellem en observation og en omkostningsfunktion i absolutte tal antages udelukkende at kunne tilskrives inefficiens.

¹² Winsten, C. B. (1957). Discussion on Mr. Farrell's Paper. Journal of the Royal Statistical, s. 282-284.

Figur 5.1 COLS vs OLS



Anm.: Illustration af COLS og OLS baseret på simuleret data. Den sorte linje markerer omkostningsfunktionen estimeret ved OLS, mens den røde linje markerer funktionen estimeret med COLS.

Kilde: Egen tilblivelse

Boks 5.1
COLS**Evaluering af metoden****Faglig robusthed**

COLS benytter en funktionel form til estimering af omkostningsfunktionen og fastsætter efficiensscorene deterministisk, derfor indeholder metoden de primære ulemperne fra både DEA og SFA. Dette gør overordnet set, at metoden er uhensigtsmæssig til benchmarking af vandsektoren, fordi efficiensscorene risikerer at give et misvisende billede af selskabernes effektiviseringspotentiale.

COLS har nogen udbredelse i den videnskabelige litteratur. Metoden anvendes dog ofte som eksempel på en uhensigtsmæssig benchmarkingmetode på grund af ovenstående problemer.

Tilgængelighed

Idet COLS er baseret på omkostningsfunktioner, vil det ikke være mere ressourcekrævende at implementere og udvikle end den SFA-model der er anvendt hidtil.

Tydelige incitamenter

Metoden er simplere end SFA, og derfor også mere intuitiv for selskaberne at forstå og agere ud fra. Selskabernes aktiviteter slår igennem 1:1 i modellen, og det er derfor nemt for selskaberne at se effekten af deres handlinger.

Metodens praktiske egenskaber

Metoden kan anvendes med de data og software som allerede anvendes i benchmarkingen i dag.

Fremadrettede anvendelsesmuligheder

COLS forventes at kunne udvikles løbende, og modellen er fleksibel i forhold til at introducere nye variable. Dog gælder for stokastiske metoder, at kravet om et tilstrækkeligt stort datasæt stiger, hvis der indføres flere variable.

Kapitel 6

Modified Ordinary Least Squares (MOLS)

Modified Ordinary Least Squares (MOLS) er en statistisk benchmarkingmetode, hvor der med OLS-regression estimeres en omkostningsfunktion, som ”modificeres” (forskydes), hvorefter alle selskaber måles op imod den forskudte omkostningsfunktion.

MOLS er første gang præsenteret i videnskabelig litteratur i 1974¹³, men har siden kun været sporadisk anvendt. I dag benyttes MOLS dog blandt andet af den danske og østrigske elnet-regulator.¹⁴

MOLS kan betragtes som en teoretisk forsimplet udgave af SFA, da SFA estimerer alle relevante parametre på samme tid, hvorimod MOLS deler processen med at beregne efficienssco-orer op i flere steps. MOLS og SFA deler derfor mange af de samme fordele og ulemper. Forsimplingen kan medføre, at det er bedre at anvende MOLS fremfor SFA i specifikke tilfælde, fx når der er få observationer i datasættet.¹⁵

Selve estimeringen af inefficiens foregår i tre trin:

1. Der estimeres en omkostningsfunktion ved brug af OLS-regression
2. Den estimerede omkostningsfunktion parallelforskydes for at tage højde for inefficiens
3. Ud fra den forskudte omkostningsfunktion kan hvert selskabs inefficiens beregnes

I det første trin estimeres en omkostningsfunktion ved brug af OLS.¹⁶ Omkostningsfunktionen fra en OLS kan fortolkes som branchens gennemsnitlige omkostninger. I en benchmarkingmodel er det dog ikke branchens gennemsnitlige omkostninger, der er interessante, men derimod branchens effektive omkostninger.

I andet trin findes derfor branchens gennemsnitlige inefficiens, som trækkes fra OLS omkostningsfunktionen. Dette vil resultere i en effektiv omkostningsfunktion og svarer til, at OLS omkostningsfunktionen parallelforskydes.

Branchens gennemsnitlige inefficiens kan beregnes ved hjælp af den såkaldte ”Method of Moments” metode.¹⁷ For at beregne den gennemsnitlige inefficiens med ”Method of Moments” op-sættes antagelser om selskabernes fordeling af inefficiens og statistisk støj. Normalt antages den statistiske støj at være normalfordelt og inefficiensen antages at være halvnormalfordelt.¹⁸

¹³ Richmond, J. (1974). Estimating the efficiency of production. *International economic review*, 515-521.

¹⁴ Den danske og østrigske regulator bruger en forsimplet version af MOLS i forhold til den, som redegøres for i gennemgangen.

¹⁵ Jf. Andor & Hesse (2014). The StoNED age: the departure into a new era of efficiency analysis? A monte carlo comparison of StoNED and the ”oldies” (SFA and DEA).

¹⁶ ordinary least squares (OLS) er standardmetoden til at estimere en regressionsanalyse

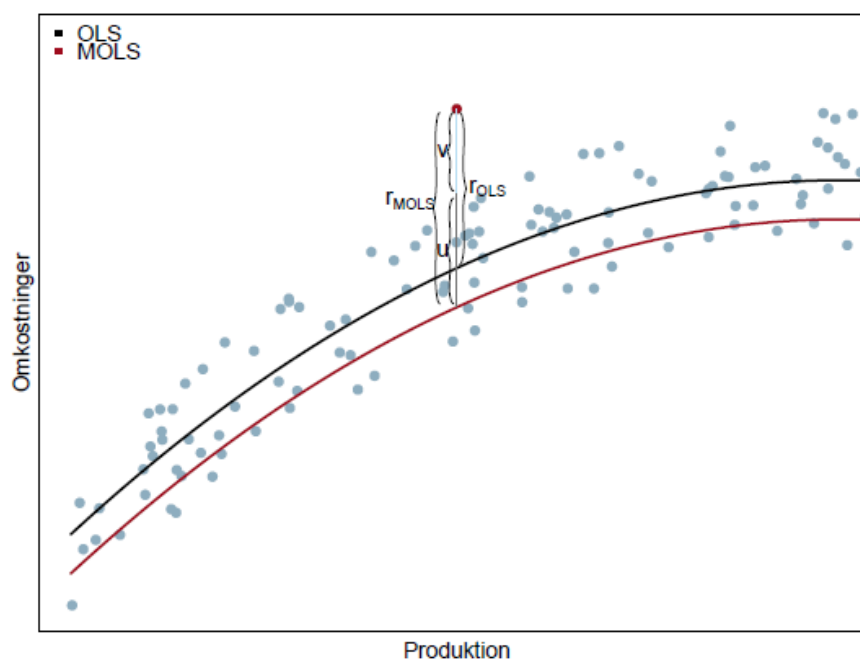
¹⁷ Method of moments er en statistisk metode til at finde en række nøgletal som fx middelværdi, varians og skævhed for en fordeling

¹⁸ Jf. Bogetoft & Otto (2010)

I det tredje, sidste trin beregnes det enkelte selskabs effektive omkostningsniveau¹⁹ ved at beregne afstanden fra selskabets faktiske omkostninger til den nye forskudte MOLS omkostningsfunktion fra trin 2, jf. Figur 6.1. Figuren viser, hvordan MOLS beregner et specifikt selskabs effektive niveau og efficiens. Den sorte linje angiver omkostningsfunktionen fra OLS (trin 1), og den røde linje angiver MOLS-omkostningsfunktionen (trin 2). Bemærk, at MOLS-omkostningsfunktionen er en parallelforskydning af OLS-omkostningsfunktionen svarende den gennemsnitlige inefficiens. Selskabets OLS-residual er angivet med ϵ_{OLS} , imens selskabets residual i MOLS er angivet med ϵ_{MOLS} .

Selskabets faktiske omkostninger fratrasket MOLS-residualet angiver selskabets effektive niveau. Den del af residualet, der stammer fra inefficiens (u) angiver selskabets efficiensscore.²⁰

Figur 6.1 **MOLS vs OLS**



Anm.: Illustration af MOLS og OLS baseret på simuleret data. Den sorte linje markerer omkostningsfunktionen estimeret ved OLS, mens den røde linje markerer funktionen estimeret med MOLS.

Kilde: Egen tilvirkning

¹⁹ Det effektive niveau angiver hvor mange omkostninger et selskab bør have givet dets output, hvis det er fuldt effektivt

²⁰ Efficiensscoren udregnes efter samme principper som i SFA, men med den forskel, at MOLS residualerne bruges fremfor SFA residualerne.

Boks 6.1 MOLS

Evaluering af metoden

Faglig robusthed

MOLS er baseret på statistisk estimering af en omkostningsfunktion, og indeholder en række af de ulemper, der almindeligvis knytter sig til fastlæggelse af funktionel form og database-grænsninger.

Beregning af efficiensscorerne tager hensyn til stokastik i data. Samtidigt har MOLS en mere enkel estimeringsproces end SFA, fordi den baserer sig på OLS-regressionen. Det giver MOLS gode teoretiske egenskaber til at estimere efficiensscorer for vandsektoren.

MOLS har til gengæld også nogle af de samme udfordringer som SFA, da modellens underliggende antagelser ikke altid er opfyldt i praksis. Eksempelvis er det ikke nødvendigvis sandt, at antagelserne om fordelingen af inefficiens og støj er korrekte. Det medfører resultater, der ikke er robuste, og som kan være misvisende.

Forsyningssekretariatet har på baggrund af de overvejende positive teoretiske egenskaber for modellen foretaget empiriske analyser for historisk data for drikkevandsselskaberne og spildevandsselskaberne. De empiriske resultater viser dog netop, at de underliggende antagelser om fordelingen mellem inefficiens og støj ikke er opfyldt når metoden anvendes til benchmarking af vandselskaberne, og at modellen derfor generelt ikke giver robuste resultater i benchmarkingen af vandselskaberne.

MOLS har lille anvendelse i videnskabelig litteratur, hvor den generelt ikke har været brugt empirisk. Det betyder, at der er begrænset kendskab og erfaringer med metodens fordele og ulempe, når den anvendes empirisk.

Tilgængelighed

Det kræver ikke yderligere ressourcer at implementere og udvikle end den nuværende SFA-model. I forbindelse med reguleringen af vandsektoren har der pågået et større arbejde med at fastlægge netop den funktionelle form. Derfor vurderes dette ikke at være en større udfordring i MOLS i forhold til den nuværende SFA-model.

Tydelige incitamenter

Vandselskabernes effektiviseringer slår ikke igennem 1:1 i metoden, eksempelvis når et selskab reducerer deres omkostninger. Det reducerer selskabernes incitament til at effektivisere, ligesom det er tilfældet med SFA-metoden.

Metodens praktiske egenskaber

Metoden kan anvendes med de data og software som allerede anvendes i benchmarkingen i dag.

Fremadrettede anvendelsesmuligheder

For at estimere effektivitet i MOLS skal der foretages en opdeling af statistik støj og inefficiens, ligesom det er tilfældet fra SFA. Vores empiriske analyser har vist, at denne opdeling kan være udfordrende. Det kan skyldes en kombination af flere elementer, herunder antallet af frihedsgrader. For få frihedsgrader øger sandsynligheden for, at antagelsen af, hvordan inefficiensen og den statistiske støj skal fordeles, bliver brudt.

Kapitel 7

Semiparametric Stochastic Frontier Analysis (Semiparametrisk SFA)

Semiparametrisk SFA er en benchmarkingmetode, som har samme tilgang til estimering af inefficiens som SFA. Det betyder, at metoden er stokastisk og skal bruge antagelser om fordelingen af støj og inefficiens. I modsætning til SFA, har semiparametrisk SFA ikke en eksplicit funktionel form. Derfor skal der ikke gøres antagelser om produktionsteknologien forud for, at inefficiensen estimeres.

Metoden blev udviklet i Fan, Li, & Weersink (1996) og ses af og til anvendt i videnskabelig litteratur.

Semiparametrisk SFA estimerer inefficiens i to trin:

1. Omkostningsfunktionen estimeres ved hjælp af en ikke-parametrisk estimator. Der findes forskellige typer af sådanne estimators, men ofte anvendes en såkaldt Kernel-regression. Kernel-regressionen har samme formål som en standard OLS regression, men anvender ikke en parametrisk funktionel form. I stedet skal man antage størrelsen på en såkaldt bandwidth som bestemmer omkostningsfunktionens udjævning.
2. Baseret på den estimerede omkostningsfunktion i trin 1, beregnes selskabernes fejllid. Fejlledene opdeles i støj og inefficiens ved hjælp af antagelser om fordelingerne af støj og inefficiens efter samme metode som i SFA. Støj og inefficiens antages ofte at være hhv. normalfordelt og positiv halvnormfordelt.

Ikke-parametriske estimators anvender forskellige metoder til at beregne funktionen. Der kan være forskel på, hvilke metoder der anvendes til forskellige estimators. Det vil være et betydeligt arbejde at undersøge dels, hvilken ikke-parametrisk estimator, som vil være mest egnet til benchmarking af vandsektoren, og derefter hvilke metoder som bør anvendes til fastsættelse af funktionen og deraf efficiensscorerne.

Hvis der anvendes en Kernel-regression, afhænger resultaterne af den valgte bandwidth. Der er mange tilgange til at fastsætte denne, men ingen konsensus om en "korrekt" metode. Andre typer af estimators har ligeledes fordele og ulemper, som for enkeltheds skyld ikke gennemgår her.

Boks 7.1 Semiparametrisk SFA

Evaluering af metoden

Faglig robusthed

Semiparametrisk SFA tager afsæt i ikke-parametrisk estimering af funktioner, hvor der ikke gøres antagelser om den overordnede funktionelle form for produktionsteknologien. Det betyder, at metoden ikke indeholder en række af de ulemper som en parametrisk metode (fx SFA) har, herunder antagelser om funktionsform. Semiparametrisk SFA gør, ligesom andre stokastiske metoder, brug af antagelser om fordelingen af støj og inefficiens. Semiparametrisk SFA fremstår fra en rent faglig vinkel som en robust metode til benchmarking af vandselskaberne. Det skyldes, at metodens store fleksibilitet er i stand til at tilpasse sig input-output-sammensætning i vandselskaberne uden at der skal gøres antagelser om produktionsteknologien.

Vores erfaring med øvrige alternativer til benchmarkingmodeller, som ligeledes anvender stokastiske metoder (MOLS, COLS og SFA), har givet indikation af, at stokastiske metoder ikke kan anvendes til at beregne robuste resultater for drikkevandsselskaber. Vi forventer derfor, at samme resultater vil være gældende for semiparametrisk SFA, hvorfor dette kriterie ikke forventes at være fuld opfyldt.

Metoden blev udviklet i 1990'erne, og har siden været anvendt sporadisk i den videnskabelige litteratur. Forsyningssekretariatet har ikke kendskab til, at metoden anvendes af andre regulatorer af forsyningsselskaber.

Tilgængelighed

Metoden er relativt avanceret og kompleksiteten i både de teoretiske og empiriske aspekter er høj. Det forventes også, at metoden vil være mere ressourcekrævende at inkludere i Forsyningssekretariatets årlige benchmarking grundet den høje kompleksitet.

Tydelige incitamenter

Vandselskabernes effektiviseringer slår ikke igennem 1:1 i metoden, eksempelvis når et selskab reducerer deres omkostninger. Det reducerer selskabernes incitament til at effektivisere, ligesom det er tilfældet med SFA-metoden.

Metodens praktiske egenskaber

Forsyningssekretariatet vil ikke kunne bruge eksisterende koder til metoden og derfor skal der foretages en større omskrivning af koderne til brug for benchmarking før en semiparametrisk SFA kan anvendes. Det kræver desuden specialiserede kompetencer at videreudvikle på modellen, fordi metoden har høj kompleksitet.

Fremadrettede anvendelsesmuligheder

For at estimere effektivitet i Semiparametrisk SFA skal der foretages en opdeling af statistik støj og inefficiens, ligesom det er tilfældet fra MOLS og SFA. I MOLS og SFA har vi empirisk vist, at denne opdeling kan være udfordrende. Det kan skyldes en kombination af flere elementer, herunder antallet af frihedsgrader. For få frihedsgrader øger sandsynligheden for, at antagelsen af, hvordan inefficiensen og den statiske støj skal fordeles, bliver brudt. Vi forventer, at de samme problemer vil være gældende for Semiparametrisk SFA, hvilket vanskeliggør de fremadrettede anvendelsesmuligheder.

Kapitel 8

Stochastic Nonlinear Envelopment of Data (StoNED)

Stochastic Nonparametric Envelopment of Data (StoNED) er en metode til benchmarking, som er udviklet tilbage i 2006.²¹ Metoden er udviklet til at fusionere DEA og SFA, dvs. den indeholder både deterministiske og stokastiske elementer. Den bruges i nogen grad i den videnskabelige litteratur og anvendes i dag af den finske regulator af elnetvirksomheder.

StoNED har tre trin til beregning af inefficiens.

1. Der estimeres en ikke-parametrisk omkostningsfunktion. Det betyder, at omkostningsfunktionen ikke antages at have en specifik funktionsform og dermed ingen parameterestimer. I stedet udgøres omkostningsfunktionen af en række lineære sektioner i forlængelse af hinanden ligesom DEA. Forskellen mellem DEA og StoNED er, at omkostningsfunktionen i StoNED ikke danner en front rundt om de mest effektive selskaber, men derimod omkring de gennemsnitlige selskaber. Det giver et residual for alle selskaber, som kan bruges til at beregne støj og inefficiens, hvilket ikke er muligt i DEA, hvor der kun beregnes inefficiens.²²
2. Residualerne opdeles i inefficiens og støj. Dette kan fx gøres ved brug af Method of Moments.²³ I StoNED bruges Method of Moments til at estimere den gennemsnitlige inefficiens, spredningen i inefficiensen samt den gennemsnitlige støj og spredningen i støjen. Det gøres ud fra antagelser om inefficiensens og støjens fordelinger. Herefter forskydes omkostningsfunktionen nedad i forhold til den gennemsnitlige inefficiens.
3. Den forskudte omkostningsfunktion fra trin 2 bruges til at estimere de minimale omkostninger for alle selskaber og derved finde efficiensscoren for hvert selskab.

Det betyder, at StoNED i trin 1 danner en front, som minder om fronten i en DEA-model, men beregner efficiensscorer i trin 2 og 3 på en måde, som minder om SFA.

Idet metoden indeholder en frontfastsættelse, som minder om den i DEA, er der de samme udfordringer med dimensionalitet²⁴ i StoNED som i en standard DEA.²⁵ Enkelte variable kan dog også indgå parametrisk i frontfastsættelse, hvilket betyder, at fronten minder mere om SFA, som bedre kan håndtere flere dimensioner.

²¹ Kuosmanen, T. (2006). Stochastic nonparametric envelopment of data: combining virtues of SFA and DEA in a unified framework.

²² Det første trin estimeres med Convex Nonparametric Least Squares (CNLS) estimator.

²³ Method of Moments er en statistisk metode, som bruges til at estimere modelparametre. Metoden er matematisk relativt mere enkelt og gør brug af en fordelings moments til estimering af modelparametre.

²⁴ Dimensionalitetsproblemet er blandt andet beskrevet i Konkurrence- og Forbrugerstyrelsen (Juni 2022) Benchmarking med forsyningssikkerhed i vandsektoren.

²⁵ Jf. Andor, M., & Hesse, F. (2014). The StoNED age: the departure into a new era of efficiency analysis? A monte carlo comparison of StoNED and the "oldies" (SFA and DEA). Journal of productivity analysis, 41(1), 85-109. s. 10.

Boks 8.1 StoNED

Evaluering af metoden

Faglig robusthed

StoNED er metodisk velfunderet, idet den er i stand til at estimere en omkostningsfunktion uden at gøre brug af antagelser om funktionel form. Samtidig håndteres stokastik i fastsættelsen af efficiensscorerne på en meningsfuld måde. Dette vurderes fordelagtigt, idet metoden ikke binder sig til antagelser om produktionsteknologien, samtidig med at den håndterer usikkerhed. StoNED vurderes derfor at estimere robuste efficiensscorer.

StoNED har de samme udfordringer som SFA, fordi metoderne fra SFA anvendes som del af StoNED-metodens steps til at beregne efficiensscore. Det betyder også, at der underliggende antagelser ikke altid er opfyldt i praksis. Eksempelvis er det ikke nødvendigvis sandt, at antagelserne om fordelingen af inefficiens og støj er korrekte, som medfører at resultater ikke er robuste og kan være misvisende.

Vores erfaring med øvrige alternativer til benchmarkingmodeller, som ligeledes anvender stokastiske metoder (MOLS, COLS og SFA), har givet indikation af, at stokastiske metoder ikke kan anvendes til at beregne robuste resultater for drikkevandsselskaber. Vi forventer derfor, at samme resultater vil være gældende for StoNED, hvorfor dette kriterie ikke forventes at være fuld opfyldt.

StoNED har nogen anvendelse i videnskabelig litteratur og anvendes i stigende grad. Den anvendes i dag af den finske elnetregulator.

Tilgængelighed

Idet StoNED har høj kompleksitet, vil metoden være mindre intuitiv for branchen, og svær at agere ud fra.

Tydelige incitamenter

StoNED giver ikke lige så stærke incitamenter til at være omkostningseffektiv, fordi selskabernes aktiviteter ikke slår igennem 1:1, eksempelvis når et selskab reducerer deres omkostninger.

Metodens praktiske egenskaber

StoNED er metodisk mere kompleks end både SFA og DEA. Derfor vil det være ressourcekrævende at udvikle en StoNED-model til brug for benchmarking. Det forventes også, at metoden vil være mere ressourcekrævende at inkludere i den årlige benchmarking grundet den høje kompleksitet.

StoNED er en beregningstung metode, der potentielt kan kræve, at der skal anvendes eksterne servere før modellen kan køres. Det afhænger af, hvor mange variable og selskaber der inkluderes i modellen. Det kræver desuden specialiserede kompetencer at videreudvikle på modellen, fordi metoden har høj kompleksitet.

Fremadrettede anvendelsesmuligheder

For at estimere effektivitet i StoNED skal der foretages en opdeling af statistik støj og inefficiens, ligesom det er tilfældet fra MOLS og SFA. I MOLS og SFA har vi empirisk vist, at denne opdeling kan være udfordrende. Det kan skyldes en kombination af flere elementer, herunder antallet af frihedsgrader. For få frihedsgrader øger sandsynligheden for, at antagelsen af, hvordan inefficiensen og den statiske støj skal fordeles, bliver brudt. Vi forventer, at de samme problemer vil være gældende for StoNED, hvilket vanskeliggør de fremadrettede anvendelsesmuligheder.

Kapitel 9

Key Performance Indicators (KPI)

Key Performance Indicators (KPI) er en udbredt metode til økonomisk performance evaluering i business management og har sit afsæt i økonomistyring. KPI er en mere enkel tilgang til økonomisk performance benchmarking, hvor der typisk opstilles ratioer mellem forskellige nøgletal, fx dækningsgrad²⁶ eller specifikke målsætninger såsom kundetilfredshed.

Analyseres dækningsgrad med KPI, vil selskabet med den højeste dækningsgrad være det mest effektive selskab og sætte benchmarket for de øvrige selskaber. Derved kan KPI bruges til at sætte mål for eksempelvis afdelinger i en virksomhed. En anden tilgang til at sætte mål ved brug af KPI er at sætte en fast målsætning for, at dækningsgraden skal være minimum 50 pct. det kommende år.

Som værktøj til målsætning for netop økonomistyring eller compliance, er KPI et velegnet værktøj. Det skyldes metodens simple tilgang til opsætning af målsætninger, samt at den er matematisk intuitiv. Netop derfor har KPI været anvendt længe, og er stadig udbredt som performanceværktøj i mange virksomheder. KPI kan bruges med det nuværende data for vandselskaberne, og vil både praktisk set og formidlingsmæssigt være relativt let at implementere i benchmarking.

Som regulatorisk benchmarkingværktøj har KPI dog en række ulemper, hvilket er beskrevet i Bogetoft og Otto (2011)²⁷. Rent fagligt er der især to væsentlige udfordringer ved metoden i forhold til benchmarking.

1. Det antages i KPI *altid*, at der er konstant skalaafkast.²⁸ Dette er ikke nødvendigvis forkert, men det lægger en væsentlig begrænsning på anvendelsen af KPI, såfremt konstant skalaafkast ikke er retvisende forhold mellem de enheder der benchmarkes på.
2. KPI et såkaldt partielt benchmarkingmål, hvor der kun analyseres på ét specifikt input-output forhold. Det betyder, at det antages, at der ikke er substitution mellem hverken inputs eller outputs. Hvis et selskab benytter to inputs, til at producere et output, vil et skift i inputsammensætningen derfor ikke blive medregnet til beregning af efficiensen. Det betyder, at selskaberne ikke fremstår mere effektivt målt på en KPI for input 2, hvis de investerer i input 2, og derved kan reducere input 1.

²⁶ Dækningsgrad betegner et selskabs overskud i procent, og beregnes som forholdet mellem overskud og omsætning. Har man fx en dækningsgrad på 50 pct., betyder det at halvdelen af omsætningen er overskud.

²⁷ Bogetoft & Otto, Benchmarking with DEA, SFA, and R, 2011

²⁸ Konstant skalaafkast betyder, at der skal anvendes samme mængde inputs til at producere en enhed output uanset størrelse. Det betyder, at der ikke er en produktionsmæssig fordel i fx at have en stor produktion. Der kan også være stigende og faldende skalaafkast.

Boks 9.1
KPI**Evaluering af metoden****Faglig robusthed**

Den største faglige udfordring til KPI er, at det er et værktøj til partielt produktivitetsmål. Det gør det svært at lave én samlet, retvisende efficiensscore for vandselskaberne. Det skyldes dels, at vandselskabernes forskellige input-output-sammensætning ikke på en retvisende måde kan sammenlignes og dels, at der skal laves flere KPI'er. Der er ikke en klar metode til at vægte de forskellige KPI'er, og derfor ikke en entydigt robust måde at få en efficiensscore ved brug af KPI.

KPI har været anvendt længe og anvendes fortsat til at sætte mål og til at følge op på strategier. Forsyningssekretariatet har ikke kendskab til, at andre regulatorer anvender KPI til økonomisk benchmarking, men metoden anvendes dog af Miljøstyrelsen i deres performance-benchmarking af vandselskaberne.

Tilgængelighed

KPI er mere enkel, og metoden vil kunne implementeres i benchmarking hurtigt, ligesom den vil være nem at udvikle.

Tydelige incitamenter

KPI giver ikke nødvendigvis selskaberne incitament til at være omkostningseffektive. Det skyldes, at de omkostningseffektiviseringer som selskaberne laver, ikke nødvendigvis slår igen i efficiensscorer, fordi KPI ikke tager hensyn til selskabernes input-outputsammensætning. Derudover skal der laves flere KPI'er, som kan påvirke hinanden. Det medfører, at et selskabs investering kan øge effektivitet målt på en KPI, men svække effektiviteten på en anden, uden at metoden retvisende kan tage højde herfor.

Metodens praktiske egenskaber

Der er ingen praktiske udfordringer med at implementere metoden. Metoden kan anvendes med de data og software som allerede anvendes i benchmarkingen i dag.

Fremadrettede anvendelsesmuligheder

Det vil være nemt at lave nye KPI'er for nye variable, såfremt der skal laves tilføjelser til benchmarkingmodellen.

Kapitel 10

Referencer

Andor, M., & Hesse, F. (2014). The StoNED age: the departure into a new era of efficiency analysis? A monte carlo comparison of StoNED and the "oldies" (SFA and DEA). *Journal of productivity analysis*, 41(1), 85-109. s. 10

Bogetoft, P., & Otto, L. (2010). *Benchmarking with DEA, SFA, and R*. Springer Science & Business Media.

Bogetoft, P., & Otto, L. (2011). *Benchmarking with DEA, SFA, and R*. Springer.

Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1), s. 1-25.

Charles, V., Aparicio, J., & Zhu, J. (2019). The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis. *European Journal of Operational Research*, s. 929-940.

Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), s. 429-444.

Da Silva, T., Martins-Filho, C., & Ribeiro, E. (2016). Comparison of nonparametric efficiency estimators: DEA, FDC, DEAC, FDCH, order-m and quantile. *Economic Bulletin*, 36(1), s. 118-131.

Daraio, C., & Simar, L. (2007). Conditional nonparametric frontier models for convex and non-convex technologies: a unifying approach. *Journal of productivity analysis*, 28(1), s. 13-32.

Dyson, R. G., & Estelle, A. S. (2010). Data envelopment analysis, operational research and uncertainty. *Journal of the Operational Research Society*, 61(1), s. 25-34.

Fan, Y., Li, Q., & Weersink, A. (1996). Semiparametric estimation of stochastic production frontier models. *Journal of Business & Economic Statistics*, 14(4), s. 460-468.

Färe, R., Grosskopf, S., & Whittaker, G. (2007). Network dea. I R. Färe, S. Grosskopf, & G. Whittaker, In *Modeling data irregularities and structural complexities in data envelopment analysis* (s. 209-240). Boston: Springer.

Hatami-Marbini, A., Emrouznejad, A., & Tavana, M. (2011). A taxonomy and review of the fuzzy data envelopment analysis literature: two decades in the making. *European journal of operational research*, 3, s. 457-472.

Konkurrence- og Forbrugerstyrelsen. (Juni 2022). Benchmarking med forsyningssikkerhed i vandsektoren. Hentet fra KFST: <https://www.kfst.dk/media/rkdn2nll/20220607-benchmarking-med-forsyningssikkerhed-i-vandsektoren.pdf>

Kuosmanen, T. (2006). Stochastic nonparametric envelopment of data: combining virtues of SFA and DEA in a unified framework.

Land, K. C., Lovell, C. A., & Thore, S. (1993). Chance-Constrained Data Envelopment Analysis. *Managerial and Decision Economics*, 14(6), s. 541-554.

Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), s. 50-60.

Olesen, O. B. (2006). Comparing and combining two approaches for chance constrained DEA. *Journal of Productivity Analysis*, 26(2), s. 103-119.

Richmond, J. (1974). Estimating the efficiency of production. *International economic review*, 515-521.

Sengupta, J. K. (1992 A). A fuzzy systems approach in data envelopment analysis. *Computers and Mathematics with Applications*, 8, s. 259-266.

Sengupta, J. K. (1992 B). Measuring efficiency by a fuzzy statistical approach. *Fuzzy Sets and Systems*, 1, s. 73-80.

Silva, T. D., Martins-Filho, C., & Ribeiro, E. (2016). A comparison of nonparametric efficiency estimators: DEA, FDC, DEAC, FDCH, order-m and quantile. *Economic Bulletin*, 36(1).

Winsten, C. B. (1957). Discussion on Mr. Farrell's Paper. *Journal of the Royal Statistical*, s. 282-284.

Teknisk gennemgang af DEA-Order-M

Teknisk bilag 2

Februar 2024



KONKURRENCE- OG FORBRUGERSTYRELSEN

Teknisk gennemgang af DEA-Order-M

Konkurrence- og Forbrugerstyrelsen

Carl Jacobsens Vej 35

2500 Valby

Tlf.: +45 41 71 50 00

E-mail: kfst@kfst.dk

Teknisk gennemgang af DEA-Order-M er udarbejdet af Konkurrence- og Forbrugerstyrelsen.

Februar 2024

Indhold

Kapitel 1	4
Indledning.....	4
Kapitel 2	5
Beskrivelse af DEA og Order-M	5
2.1 Beskrivelse af DEA	5
2.2 Beskrivelse af Order-M	6
Kapitel 3	10
Fastsættelse af parameteren M.....	10
3.1 Bud på fastsættelse af egenskaber for at vælge M og analyse af egenskaber	11
3.2 Bagvedliggende analyser af egenskaber for M	12
Kapitel 4	23
Opsamling.....	23
Appendix A	24
Matematisk formulering af DEA og DEA-Order-M	24
Referencer	26

Kapitel 1

Indledning

I dette tekniske baggrundsnotat redegøres der for metoden DEA-Order-M (herefter Order-M), ligesom der redegøres for de egenskaber og analyser, som kan anvendes i forbindelse med fastsætter parameteren "M" i metoden. Der tages udgangspunkt i de empiriske resultater fra benchmarkingen af drikkevandsselskaberne.

Order-M er en simpel udvidelse af den standard DEA-metode (DEA), som anvendes i benchmarkingen af vandselskaberne i dag. Der tages derfor udgangspunkt i DEA i beskrivelsen af Order-M. Order-M udjævner fronten i forhold til DEA, således at flere selskaber end de efficiente også påvirker fronten. Det gør fronten mindre afhængig af de efficiente selskaber og mere lempelig end i en standard DEA. Det betyder, at Order-M altid stiller selskaberne bedre end i en standard DEA.

Forsyningssekretariatet vurderer, at Order-M er en mere retvisende model, end den kombination af DEA og SFA, som anvendes til totaløkonomisk benchmarking af vandselskaberne i dag. Order-M giver mere retvisende, robuste og gennemsigtige resultater, som indebærer mindre usikkerhed. Det er både til gavn for vandselskaberne og forbrugerne.

Kapitel 2

Beskrivelse af DEA og Order-M

2.1 Beskrivelse af DEA

Data Envelopment Analysis (DEA)¹ er en matematisk metode, der anvender lineær programmering til at sammenligne en række enheder – i denne sammenhæng en gruppe af vandselskaber med henblik på at bestemme effektiviteten for hvert selskab.

Selskaberne sammenlignes på baggrund af deres output holdt op mod deres input. I regulering af vandselskaberne bruger Forsyningssekretariatet selskabernes omkostninger som input og de såkaldte OPEX- og CAPEX-netvolumenmål som outputs. Netvolumenmålene er en aggregering af flere underliggende outputs, hvor OPEX er et udtryk for de operationelle outputs (fx produktion af rent drikkevand eller rensning af spildevand) og CAPEX er et udtryk for de outputs, der relaterer sig til selskabernes anlægsaktiver (fx de km ledning et selskab ejer).

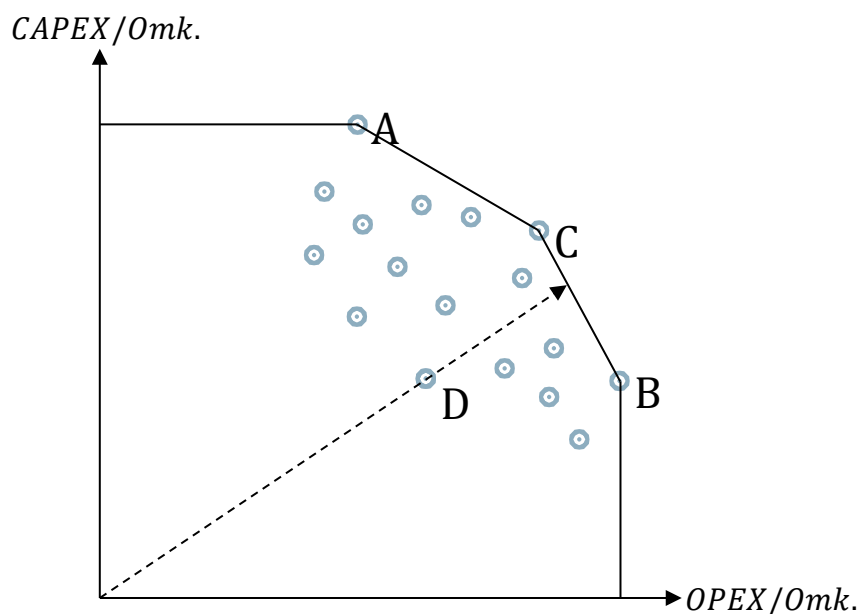
Figur 2.1 viser en benchmarkingmodel med fire selskaber (A, B, C og D). DEA fastsætter en front, som udgøres af de mest effektive selskaber; A, C, B (frontselskaberne). Selskab A er det selskab, der producerer mest CAPEX i forhold til sine omkostninger, og det betegnes derfor som værende efficient (fuldt effektiv). Selskab B er det selskab, der producerer mest OPEX i forhold til sine omkostninger, og det betegnes derfor også som værende efficient. Selskab C betegnes også som værende efficient. Det skyldes, at det, givet en bestemt vægtning mellem CAPEX og OPEX, er det selskab, der producerer mest i forhold til sine omkostninger.

DEA-fronten beregnes i eksemplet i figur 2.1 på baggrund af de tre efficiente selskaber og udgøres af de fire fuldt optrukne streger i figuren. Selskab D ligger under fronten og er derfor inefficient. Afstanden til fronten (angivet af den stiplede linje) afgør selskab D's efficiensscore. En efficiensscore kan ligge mellem 0 og 100 pct. Hvis efficiensscoren fx er 80 pct., betyder det, at et selskab kan reducere sine omkostninger til 80 pct. af de faktiske omkostninger og fortsat producere den samme mængde OPEX og CAPEX.²

¹ Charnes, Cooper, & Rhodes (1978)

² Det antages her, at modellen er input-orienteret. Hvis den derimod var output-orienteret defineres en efficiensscore til, hvor meget output kan øges for et fast niveau af input.

Figur 2.1 Illustrativt eksempel af DEA



Anm.: Figuren viser en DEA model med to outputs og ét input. De sorte optrukne linjer indikerer fronten som det ineffektive selskab (D) måles op imod.

Kilde: Egen tilvirkning

Den matematiske formulering af DEA er angivet i Appendix A.

2.2 Beskrivelse af Order-M

Order-M er en simpel udvidelse til en standard DEA, som kan siges at udjævne fronten, således at fronten dannes på baggrund af flere selskaber og ikke kun de efficiente selskaber. Fordelen ved at udjævne fronten er, at Order-M bliver mindre afhængig af frontelskaberne fra en standard DEA. Det betyder, at Order-M-fronten bliver repræsenteret af flere selskaber og på den måde tager højde for forskellige former for støj som fx datausikkerhed, generelle udsving, og risiko for at frontelskaberne ikke er fuldt repræsentative for resten af sektoren.

Order-M blev første gang foreslået i 2002³ til at analysere outliers i DEA. Metoden er efterfølgende hyppigt anvendt som benchmarkingmetode til analyser af forskellige sektorer.⁴

Order-M udjævner fronten ved at beregne DEA flere gange – altså med flere iterationer. I hver iteration udtrækkes M-antal tilfældige selskaber, som potentielt kan være med til at danne fronten.⁵ De mest effektive selskaber ud af M-selskaberne, vil blive frontelskaber i den givne

³ Cazals, Florens, & Simar (2002)

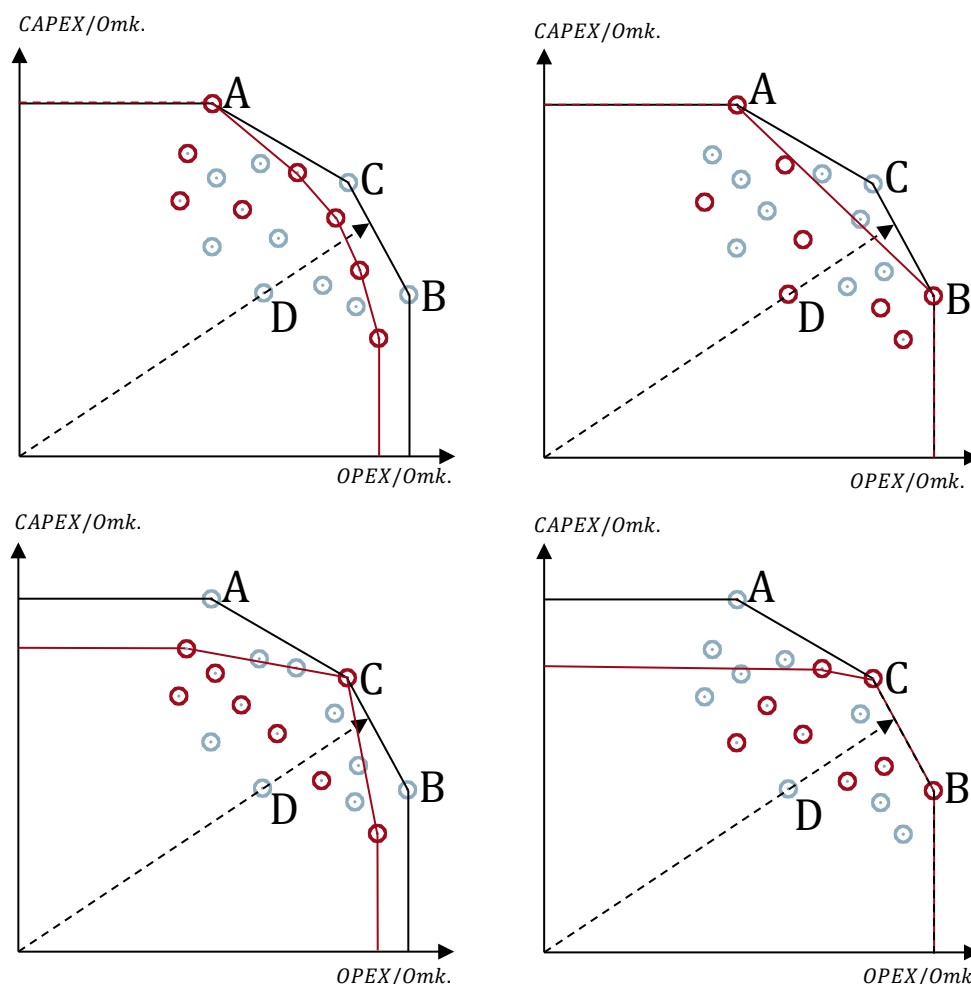
⁴ Silva, Martins-Filho, & Ribeiro (2016) og Daraio & Simar (2007)

⁵ Selskaberne trækkes med tilbagelæsning, således at ét selskab kan trækkes flere gange.

iteration. Dette adskiller sig fra DEA, hvor det er de efficiente selskaber ud af alle selskaber, der bliver frontelskaber. Hvis de efficiente selskaber ikke bliver udtrukket i Order-M, vil det blive de næstmest effektive selskaber, der udgør fronten i den givne iteration. Den endelige Order-M-efficiensscore beregnes som den gennemsnitlige efficiensscore over alle iterationerne. Det betyder, at Order-M altid giver højere efficiensscore end DEA.

Figur 2.2 illustrer, hvordan fronten i fire tilfældige Order-M-iterationer bliver påvirket af, hvilke selskaber der indgår i udtrækket. Figuren viser samtidig samme DEA-front som Figur 2.1. De røde punkter indikerer de selskaber, der er tilfældigt udtrukket, og de blå punkter indikerer de selskaber, som ikke er blevet udtrukket i de konkrete iterationer. I øverste figur til venstre er selskab C og B fx ikke udtrukket. Det vil sige, at selskab C og B ikke kan være med til at danne fronten i den givne iteration. Dermed dannes fronten (angivet ved de røde linjestykker) på baggrund af selskab A samt de næstmest effektive af de røde selskaber. Det betyder, at de inefficiente selskaber nu sammenlignes med den røde front i stedet for den sorte, hvilket gør, at de ligger tættere på fronten og dermed får en højere efficiensscore. Bemærk, at fronten i hver iteration kan variere meget afhængig af, hvilke selskaber der indgår i udtrækket. I nogle iterationer vil fronten være identisk med DEA-fronten. Dette sker, hvis både selskab A, B og C trækkes i den enkelte iteration. Fronten kan dog aldrig gå nordøst for DEA-fronten. Derfor er den gennemsnitlige efficiens over alle iterationerne i Order-M altid højere end i en standard DEA.

Figur 2.2 Sammenligning af en DEA-front og fronten fra enkelte Order-M-iterationer

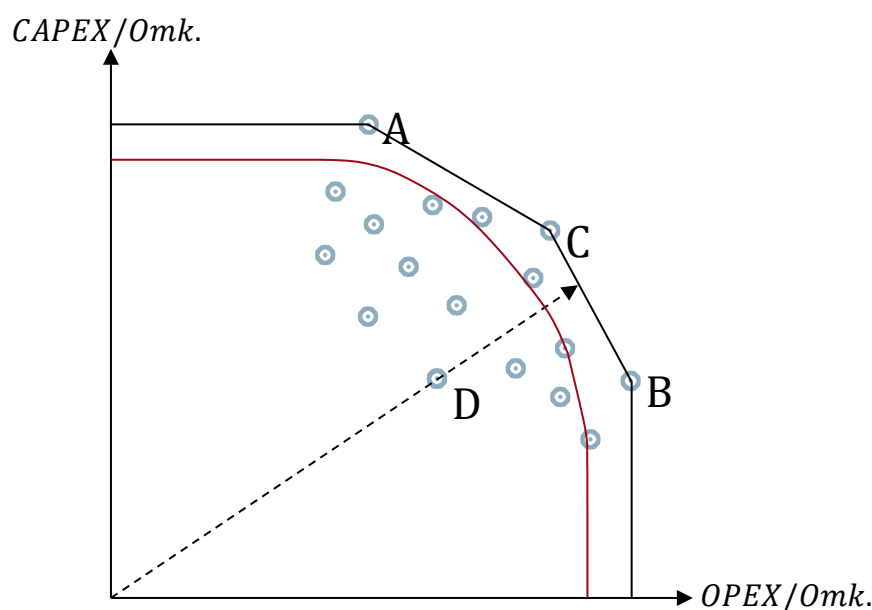


Anm.: Figuren sammenligner fronten i forskellige Order-M-iterationer med en standard DEA. Hvert punkt indikerer et selskab, hvor de røde punkter er de selskaber, der er tilfældigt udtrukket til at må indgå på fronten i den givne iteration, og de blå punkter er de selskaber, der ikke indgår i udtrækket. De sorte optrukne linjer angiver DEA-fronten, og de røde linjer angiver fronten i den givne Order-M-iteration. Figuren er lavet på simuleret data.

Kilde: Egen tilvirkning

At tage et gennemsnit af efficiensscorene fra alle iterationerne svarer til, at den endelige Order-M-front beregnes som et gennemsnit af fronterne fra hver iteration, jf. Figur 2.3. Figuren viser, hvordan Order-M-fronten ser ud i dette eksempel, når der tages et gennemsnit af fronterne fra alle underlæggende iterationer. Det ses, at Order-M-fronten ligger sydvest fra DEA-fronten, hvilket gør den mere lempelig overfor selskaberne end fronten fra en standard DEA. Det betyder, at selskaberne får en højere efficiensscore og dermed et lavere individuelt effektiviseringskrav end ved brug af en standard DEA-model. Figuren viser samtidig, hvordan enkelte selskaber kan ligge nordøst for Order-M-fronten, hvilket betyder, at de får en efficiensscore over 1. Det skyldes, at Order-M tager højde for støj, som lader selskaberne fremstå mere effektive, end hvad der ifølge modellen som udgangspunkt er teknisk muligt.

Figur 2.3 Sammenligning af fronten i DEA og Order-M



Anm.: Figuren sammenligner fronten i DEA og Order-M. Hvert punkt indikerer et selskab. De sorte optrukne linjer angiver DEA-fronten, og de røde linjer angiver Order-M-fronten. Figuren er lavet på simuleret data.

Kilde: Egen tilvirkning

Order-M bruger det samme data og antagelser som DEA. Det er derfor kun nødvendigt at tage stilling til to ekstra parametre, ved anvendelse af Order-M som benchmarkingmetode.

- 1) Antallet af iterationer: Antallet af iterationer skal sættes så højt som muligt, således at resultaterne ikke varierer, hvis modellen genberegnes senere. Order-M kræver dog meget computerkræft, hvorfor antallet af iterationer er en balance mellem, hvornår resultaterne konvergerer, og hvor meget tid det må tage at genberegne modellen. For Forsyningssekretariatets konkrete model viser analyserne, at modellen har konvergeret ved 5.000 iterationer, og at der derfor ikke er behov for yderligere iterationer. Forsyningssekretariatet anvender derfor 5.000 iterationer.
- 2) Antal træk (M): M angiver, hvor mange selskaber der trækkes i hver iteration.⁶ Jo lavere M er, desto mere udjævnes fronten. Det skyldes, at et lavt M øger sandsynligheden for, at de efficiente selskaber ikke bliver udtrukket i en given iteration. Fastsættelsen af M diskuteres og analyseres i kapitel 3.

Den matematiske formulering af Order-M er vist i Appendix A.

⁶ Bemærk, at Forsyningssekretariatet trækker selskaber med tilbagelægning. Det vil sige, at det samme selskab kan trækkes flere gange, hvilket medfører, at der i alt trækkes færre selskaber end M tilsiger.

Kapitel 3

Fastsættelse af parameteren M

Ved anvendelse af en Order-M-model skal parameteren M fastsættes for at bestemme, hvor mange selskaber, der skal trækkes i hver iteration. Forsyningssekretariatet har ikke kendskab til entydige metoder til at fastsætte M. Derfor giver Forsyningssekretariatet i dette kapitel selv et bud på en metode til fastsættelse af et fagligt og robust M.

For at fastsætte parameteren M har Forsyningssekretariatet opstillet et bud på en række egenskaber, som Order-M bør leve op til. På den måde kan der vælges det M, der medfører, at egenskaberne i højest muligt omfang bliver opfyldt. Egenskaberne har tilknytning til de evalueringskriterier, som er Forsyningssekretariatet bruger til at evaluere forskellige alternativer til benchmarkingmetode. Evalueringskriterierne er beskrevet i *Ændringer til metode for benchmarking*.

Egenskaberne er opstillet med henblik på varetagelse af flere hensyn. En af hovedårsagerne til at bruge Order-M fremfor DEA, at benchmarkingen bør være mindre følsom overfor de efficiente selskaber. På den måde vil enkeltsselskaber ikke have væsentlig betydning for resultaterne for de øvrige selskaber. Derudover bør resultaterne af benchmarkingen bidrage til, at vandselskaberne har incitament til at drive vandselskabet effektivt. Her er et realistisk estimat af selskabernes effektiviseringspotentiale væsentligt.

Med henblik på at fastsætte M har Forsyningssekretariatet derfor opstillet følgende bud på egenskaber for Order-M, som bør kunne opfyldes ved fastsættelse af M:

- » Resultater er ikke følsomme over for data fra enkelte selskaber
- » Resultater tager hensyn til støj uden at udvande formålet om identificering af efficiens
- » Fronten udgøres af mange selskaber, så fronten har høj en grad af fleksibilitet
- » Modellen opfylder betingelser og antagelser fra DEA

Resultater er ikke følsomme over for data fra enkelte selskaber

Resultatet af benchmarkingmodellen bør ikke være følsom over for, om der er ukorrekte data for et af frontsselskaberne. Denne egenskab skal bidrage til, at der ikke er et enkelt selskab, der er definerende for den standard, som alle selskaber måles op imod. Egenskaben indebærer, at benchmarkingen hverken er følsom overfor om enkelte frontsselskaber er repræsentative for branchen eller ej, eller om data for enkelte frontsselskaber er forkerte.

Resultater tager hensyn til støj uden at udvande formålet om identificering af efficiens

En af svaghederne ved en standard DEA-model er, at DEA ikke kan håndtere støj. Denne svagthed bliver håndteret i Order-M, men håndteringen i metoden kan være på bekostning af målt efficiens. Når der sættes et lavere M, indføres der samtidig både mere støj og højere målt efficiens. Egenskaben indebærer, at fastsættelse af M ikke reduceres til et niveau, hvor identificering af inefficiens udvandes.

Fronten udgøres af mange selskaber, så fronten har en høj grad fleksibilitet

I anvendelsen af en standard DEA-model i forbindelse med Forsyningssekretariatets benchmarking af vandselskaberne er der typisk kun 3-4 frontskaber hvert år.⁷ Dette gør, at modellen ikke er fleksibel, da enkelte selskaber således i relativ stor grad påvirker resultaterne for de øvrige selskaber. En af styrkerne ved Order-M er, at de samme 3-4 frontskaber kun indgår delvist i Order-M-fronten. Det skyldes, at disse frontskaber ikke bliver udtrukket i alle iterationerne i Order-M. Derfor danner andre selskaber front i disse iterationer. Dermed indgår der flere selskaber i Order-M-fronten sammenlignet med en standard DEA-model, og derfor er fronten mere fleksibel. Egenskaben indebærer, at fronten udgøres af mange selskaber, hvilket understøtter fleksibilitet og dermed robusthed.

Modellen opfylder betingelser og antagelser fra DEA

Som nævnt tidligere er Order-M en udbygning af en standard DEA-model, hvor Order-M består af en gennemsnitsbetragtning af flere tusinde iterationer af en DEA-model. Derfor er det også centralt, at de antagelser og forudsætninger, der gøres for den overordnede Order-M-model, fører til retvisende resultater i hver iteration af en DEA-model. Hver enkel iteration i Order-M skal derfor også leve op til de generelle forudsætninger, der bør være opfyldt i en standard-DEA. En af forudsætningerne for en retvisende DEA-model er, at der skal være et minimum af observationer for at resultaterne er brugbare.

3.1 Bud på fastsættelse af egenskaber for at vælge M og analyse af egenskaber

Valget af M påvirker, i hvilken grad ovenstående egenskaber opfyldes. M bør fastsættes, så flest mulige egenskaber bliver opfyldt i højest mulig grad. Denne fremgangsmåde vil ikke give et endeligt, præcist M, men vil hjælpe med at indsnævre et spænd for, hvilke M'er der bør anvendes i en Order-M-model.

Til brug for beregning af effektiviseringskrav for vandselskaberne, vil den endelige beslutning om et M skulle være et enkelt tal og ikke et spænd. Derfor skal det afgøres, hvor i spændet for M den endelige model skal estimere effektivitet. Denne fastsættelse gennemgås ikke i nærværende notat, da den vil afhænge af det konkrete data, der ligger til grund for modellen.

Forsyningssekretariatet har foretaget analyser og undersøgelser, som indikerer, at det er muligt at indsnævre fastsættelsen af M til et spænd og samtidig opfylde egenskaberne. Egenskaber, analysemetode og umiddelbare konsekvenser for valg af M er opsummeret i Tabel 3.1.

Resultaterne i Tabel 3.1 er alle baseret på det data, som blev indberettet af drikkevandsselskaberne til brug for benchmarking og effektiviseringskrav til økonomiske rammer for 2023-2024. Forsyningssekretariatet har også foretaget analyser ved brug af historisk data for både drikkevandsselskaberne og spildevandsselskaber, som overordnet viser de samme resultater. Disse resultater er derfor udeladt i dette notat. Analyserne fokuserer primært på drikkevandsselskaberne, da det er drikkevandsselskaberne, der skal benchmarkes i den næstkommende benchmarking. I alle analyserne er der blevet fjernet én outlier fra frontfastsættelsen, så der indgår 73 selskaber.⁸

⁷ Antallet af frontskaber i en standard DEA-model afhænger i høj grad af, hvor mange input og output der indgår i modellen, samt hvor forskelligartede selskaberne er. I vores DEA-model anvender vi kun 1 input og 2 output, hvilket gør, at antallet af frontskaber typisk vil være lavt.

⁸ Det selskab, der er identificeret som en outlier, får stadig tildelt et resultat. En outlier betyder blot, at selskabet aldrig kan indgå i frontfastsættelsen. Selskabet er fjernet på baggrund af en høj superefficiens, der betyder, at selskabets efficiensscore afviger i så væsentlig grad fra de øvrige selskaber, at det er usandsynligt, at andre selskaber kan sammenlignes med selskabet.

Såfremt Order-M implementeres forventes det, at der bruges de samme analyser til benchmarkingen af drikkevandsselskaberne i 2025 og til benchmarkingen af spildevandsselskaberne i 2026.

Tabel 3.1 Opsummering af egenskaber, analysemetode og konsekvenser for valg af M

Egenskab	Analysemetode	Umiddelbar konsekvens for størrelsen af M i benchmarkingen for 2023
Resultater er ikke følsomme over for data fra enkelte selskaber	Måle betydningen af at fjerne selskaber fra modellen	Opfyldt for alle M
Resultater tager hensyn til støj uden at udvande målt efficiens	Måle hvor stor ændringen er til den gennemsnitlige efficiensscore for hele sektoren	Minimum 30
	Beregne om der er signifikant forskel mellem DEA og Order-M-resultater ved brug af Mann-Whitney-U-test	Cirka 24
Fronten fra Order-M har høj fleksibilitet	Tælle hvor ofte et selskab udtrækkes og danner front i iterationerne	Maximum 50
Modellen opfylder betingelser og antagelser fra DEA	Tælle antallet af selskaber, der placerer sig over fronten og dermed identificeres som værende superefficiente	Minimum 28
	Evaluere minimumsbetingelser for en DEA, særligt med henblik på minimumsantallet af observationer i en DEA-model	Minimum 13

Kilde: Egen tilvirkning

3.2 Bagvedliggende analyser af egenskaber for M

1. Resultater er ikke følsomme over for data fra enkelte selskaber

Sådan analyseres egenskaben

Denne egenskab analyseres ved at undersøge ændringerne på branchens efficiensscorer, når der fjernes ét specifikt selskab fra modellen. Analysen gentages flere gange, hvor der fjernes et nyt specifikt selskab hver gang. Analysen kan derfor også indikere, hvorvidt modellen er robust over for ændringer i data for enkeltsselskaber.

Det viser resultaterne

Resultaterne af analysen viser, at Order-M-resultaterne ikke er følsomme over for data for enkelte selskaber, uanset hvilket M der vælges. Resultaterne viser derfor, at der kan vælges et relativt bredt spænd for M, hvor denne egenskab stadig vil være opfyldt. Det betyder dog samtidig, at resultaterne fra denne analyse ikke medvirker til at indsnævre valget af M i særlig høj grad.

Resultaterne er herudover med til at understøtte, at Order-M-modellen generelt er robust over for betydningen af enkelte selskaber, hvilket er positivt og med til at bekræfte, at Order-M lever op til de generelle evalueringskriterier, Forsyningssekretariatet har sat op for valg af ny benchmarking-model.

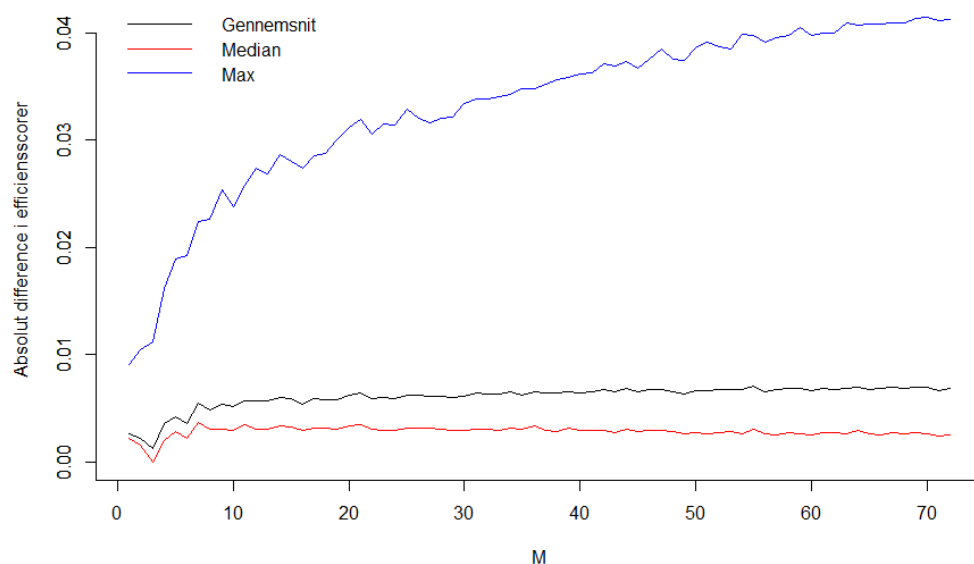
Uddybning af den specifikke analyse for egenskaben

Som tidligere nævnt er Order-M med til at udjævne fronten, så resultaterne bliver mindre følsomme overfor de efficiente selskaber i forhold til det, der er tilfældet i en standard DEA-

model. Det medfører, at de næstmest effektive selskaber begynder at påvirke fronten i stedet. På den måde repræsenterer fronten i Order-M flere selskaber end ved en standard DEA-model, men i Order-M vil hvert selskab vægte mindre i estimatet af, hvor og hvordan fronten er placeret i Order-M. Resultaterne fra Order-M bør derfor ikke være væsentligt påvirket af, hvorvidt et enkelt selskab er med i frontfastlæggelsen eller ej. Det betyder samtidig, at Order-M heller ikke bør være væsentligt påvirket af, at der sker en ændring i data for et bestemt selskab.

For at teste om enkelte selskaber påvirker resultaterne i en væsentlig grad, har Forsyningssekretariatet analyseret på konsekvenserne ved helt at fjerne det selskab, som har den største betydning for fronten. Der er i analysen fjernet et selskab danner front i en standard DEA, og hvor der samtidig er en stor andel af de øvrige selskaber i benchmarkingen, som måles op imod det konkrete selskabs performance.

Figur 3.1 Forskel i efficiensscoren mellem Order-M med alle selskaber og Order-M uden det mest indflydelsesrige selskab



Anm.: Figuren viser forskellen i efficiensscoren i en Order-M-model, hvor alle selskaber kan udtrækkes til at danne fronten og en Order-M-model, hvor det mest indflydelsesrige selskab ikke kan være med i udtrækket i nogle iterationer. Den sorte linje angiver ændringen i efficiensscoren for det gennemsnitlige selskab, den røde linje angiver medianen, og den blå linje angiver den maksimale forskel i efficiensscoren mellem de to Order-M-modeller.

Kilde: Egen tilvirkning

Figur 3.1 viser, at modellen ikke er følsom overfor enkelte selskaber uanset valget af M. Både den gennemsnitlige forskel i efficiensscoren samt medianen er forholdsvis konstant for de fleste M'er på cirka henholdsvis 0,65 procentpoint og 0,3 procentpoint (indikeret ved den sorte og røde linje).

For enkelte selskaber ændrer efficiensscoren sig med op til cirka 4 procentpoint (indikeret ved den blå linje for høje M'er). Selvom dette kan have en vis økonomisk betydning for det enkelte selskab, er ændringen stadig inden for, hvad der umiddelbart vurderes acceptabelt af påvirkning fra enkelte selskaber. Når M falder, falder påvirkningen også for dette selskab. Det

skyldes, at det indflydelsesrige selskab påvirker den oprindelige Order-M-model mindre når M falder, og at forskellen mellem de to modeller derfor også bliver mindre.

2. Resultater tager hensyn til støj uden at udvande formålet om identificering af efficiens

Sådan analyseres egenskaben

Denne egenskab analyseres ved at undersøge, hvordan valget af M påvirker branchens gennemsnitlige efficiensscore. Når M reduceres, stiger selskabernes efficiensscore, og modellen tager implicit højde for mere støj. Det er dog på bekostning af, at der er risiko for, at det udvander målet om efficiens.

Påvirkningen på branchens efficiensscore ved forskellige størrelser af M undersøges i to underanalyser:

- » Hvor stor er den marginale ændring i efficiensscoren?
- » Hvornår er der signifikant forskel i efficiensscoren mellem en standard DEA og Order-M?

Det viser resultaterne

Analysen giver ikke et specifikt, fast M, men derimod en indikation af hvornår hensynet til støj begynder at udvande målet for efficiens. Resultaterne viser, at M bør fastsættes til minimum ca. 30.

Analysen viser, at den marginale ændring i efficiensscoren ved M omkring 30 er 0,1 procentpoint. Det betyder, at når M stiger til 31, falder den gennemsnitlige efficiensscore med 0,1 procentpoint. Samtidig viser analysen af den relative marginale ændring i efficiensscorene, at måden efficiensscorene ændrer sig på ved forskellige størrelser for M, finder et stabilt niveau ved M lig med 30. Det er også ved M lig med 30, at efficiensscorene fra Order-M bliver signifikant højere end i en standard DEA. Forsyningssekretariatet vurderer, at et lavere M – og dermed større absolut marginal ændring – vil udvande målet for efficiens for meget.

Uddybning af den specifikke analyse for egenskaben

I en standard DEA antages det, at der ikke er støj i modellen. Det vil sige, at alle selskaber er repræsentative til at danne fronten, og at der ikke er støj i data. I Order-M lempes der på denne antagelse, hvilket kan fortolkes til, at en del af Order-M-fronten er udgjort af støj.

Der findes ikke et generelt mål for, hvor meget støj der skal tillades i Order-M. Det er vigtigt at fjerne nok støj til, at selskaberne ikke får urealistiske høje effektiviseringskrav. Samtidig må der ikke antages så meget støj, at modellen ikke længere kan identificere sektorens inefficiens. Vi skal derfor vælge et M, som giver en passende balance mellem støj og inefficiens.

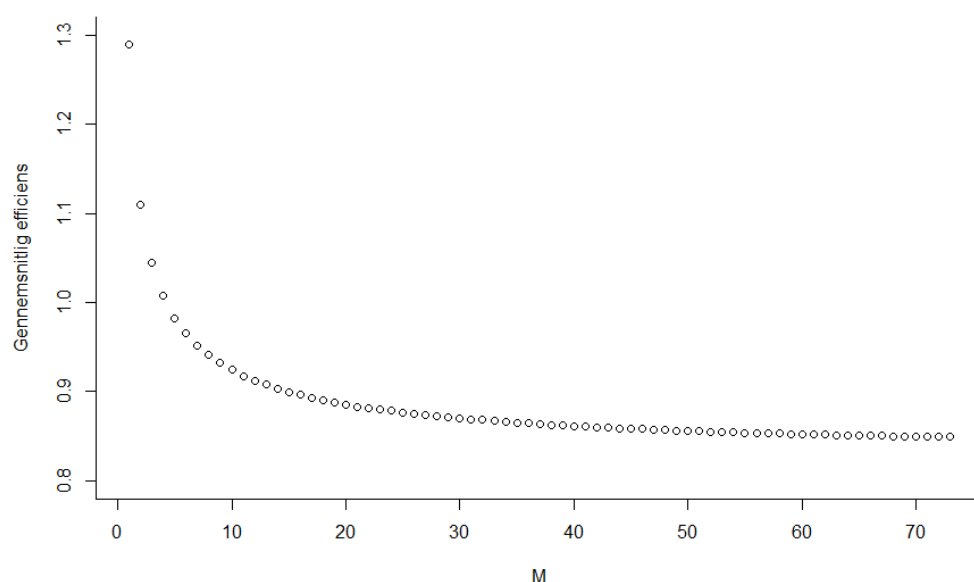
i. Marginal ændring i efficiensscore

Sammenhængen mellem støj og efficiens kan ses i Figur 3.2. Figuren viser branchens gennemsnitlige efficiens for forskellige værdier af M. Når M falder antages det implicit, at der er mere støj, hvilket resulterer i højere efficiensscore.

Figuren viser fx, at den gennemsnitlige efficiensscore er større end 1 for små værdier af M. Det betyder, at modellen tager højde for støj i så stor grad, at det gennemsnitlige selskab er mere effektivt, end hvad der burde være muligt. Her er det derfor tydeligt, at Order-M antager, at der er for meget støj. Når M stiger, konvergerer efficiensscorene mod en standard DEA, hvor

der antages ikke at være støj. Det betyder, at Order-M ikke tager meget højde for støj, når M er høj.⁹

Figur 3.2 Branchens gennemsnitlige efficiens



Anm.: Figuren viser branchens gennemsnitlige efficiensscore. En efficiensscore på 1 betyder, at branchen identificeres som værende fuldt effektiv

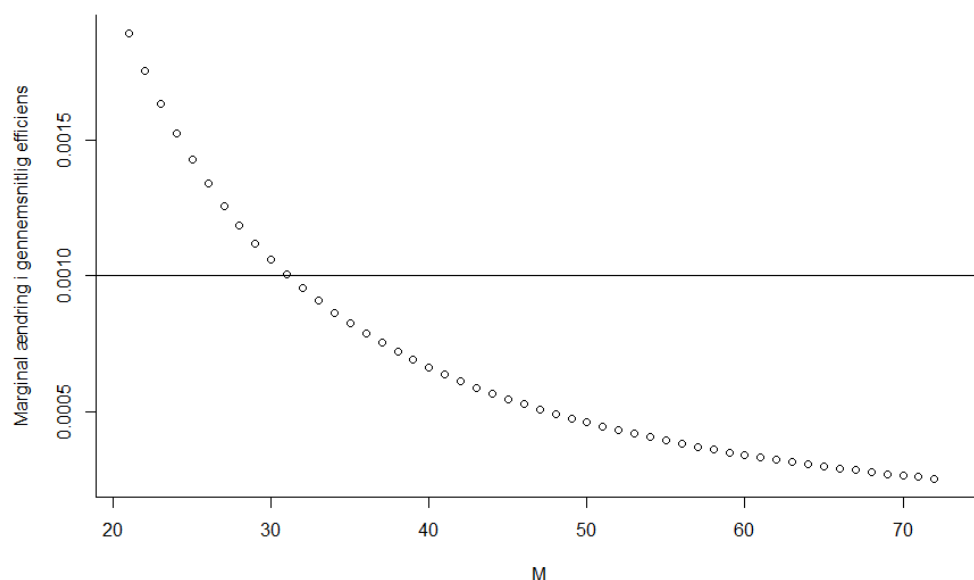
Kilde: Egen tilvirkning

Der bør findes et cut-off for, hvornår den gennemsnitlige efficiens stiger for meget, i forhold til hvor meget mere støj der tilføjes. Figur 3.3 viser den marginale ændring i den gennemsnitlige efficiens.¹⁰ Figuren viser fx, at den marginale ændring i den gennemsnitlige efficiens er under 0,05 procentpoint for høje værdier af M. Det betyder, at en marginal reduktion i M (svarende til en marginal stigning i antaget støj) øger sektorens gennemsnitlige efficiens med cirka 0,05 procentpoint for høje værdier af M.

⁹ Bemærk, at der tillades tilbagelægning, og at Order-M derfor altid vil give højere efficiensscore end DEA, selvom M er sat til det maksimale (antallet af selskaber).

¹⁰ Grafen udlines ved hjælp af $y = \frac{a}{M^b} + c$, hvor parametrene a, b og c estimeres i en regressionsmodel. Hældningen på grafen svarer derfor til hældningen på tangenten for y i hvert punkt for hele værdier af M.

Figur 3.3 Marginal ændring i branchens gennemsnitlige efficiens



Anm.: Figuren viser den marginale ændring i branchens gennemsnitlige efficiens, når M stiger med 1. Den horisontale linje indikerer Forsyningssekretariatets vurdering af en cut-off-værdi for, hvornår M ikke bør blive lavere.

Kilde: Egen tilvirkning

Forsyningssekretariatet vurderer, at en cut-off-værdi på 0,1 (markeret med den horisontale linje i Figur 3.3) er en rimelig grænse for, hvornår der antages for meget støj i modellen. Det betyder, at der her sker en forholdsvis lille ændring i efficiensscorerne ved en ændring i M.

På baggrund af ovenstående cut-off-værdi, viser første underanalyse, at M ca. bør fastsættes til minimum 30.

ii. Signifikantest for forskel mellem DEA og Order-M

Balancen mellem støj og inefficiens kan også defineres ud fra en signifikans-tankegang. Her ønsker vi, at Order-M skal give selskaberne en signifikant bedre efficiensscore i forhold til standard DEA. Resultaterne skal imidlertid ikke være mere signifikante end nødvendigt, da det igen vil udvande efficiensen.

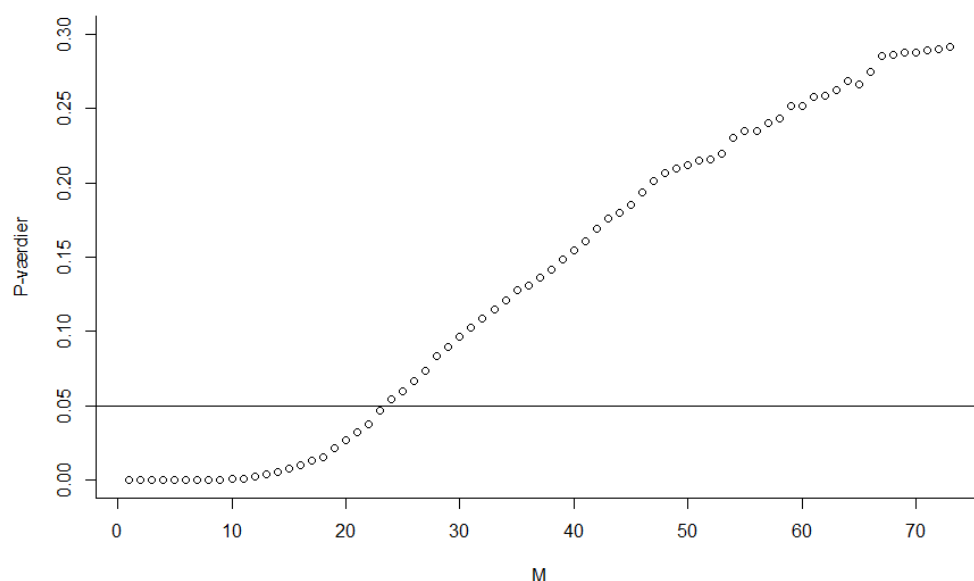
Hvorvidt Order-M efficiensscorerne er signifikant højere end DEA, måles ved brug af en såkaldt Mann-Whitney U test (herefter MW) (Mann & Whitney, 1947). MW tester sandsynligheden for, at en tilfældig efficiensscore fra Order-M-modellen er større end en tilfældig

efficiensscore fra DEA-modellen.¹¹ Når sandsynligheden er over 95 pct., siges forskellen mellem de to modeller at være signifikant.¹²

Resultaterne fra MW-testen er vist i Figur 3.4. Figuren viser, at der er en signifikant forskel mellem Order-M og DEA, når M er 24 i 2023.

Dermed viser anden underanalyse, at M cirka bør fastsættes til omkring 24.

Figur 3.4 Signifikanstest på forskellen mellem efficiensscorene i DEA og Order-M



Anm.: Figuren viser om der er signifikant forskel på efficiensscorene fra DEA og Order-M. P-værdierne angiver sandsynligheden for, at en tilfældig efficiensscore fra Order-M ikke er større end eller lig med en tilfældig efficiensscore fra DEA.

Kilde: Egen tilvirkning

3. Fronten udgøres af mange selskaber, så fronten har høj en grad af fleksibilitet

Sådan analyseres egenskaben

Egenskaben analyseres ved at undersøge, i hvor mange iterationer hvert selskab indgår på fronten. Der sættes en øvre grænse på 50 pct., så ingen selskaber må indgå på fronten i mere end 50 pct. af iterationerne. På den måde sikres det, at ingen selskaber påvirker fronten "for

¹¹ Man vil normalvis parre efficiensscorene mellem de to modeller, så man holder styr på, hvilke efficiensscore i de to modeller, der kommer fra samme selskab. Dette vil dog per definition altid give signifikante forskelle mellem Order-M og DEA. Derfor anvendes her ikke-parrede test.

¹² Der anvendes en "ensidede" alternativ hypotese.

meget". Samtidig giver det fronten en høj grad af fleksibilitet, da det tillader flere selskaber at få indflydelse på fronten.

Det viser resultaterne

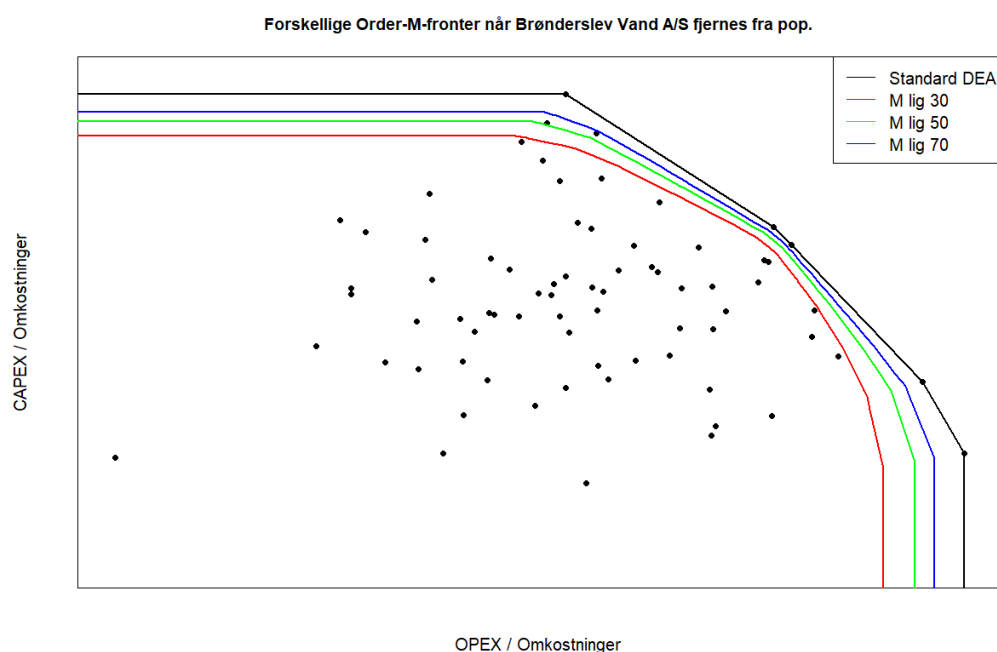
Resultaterne viser, at M maksimalt bør være lig med 50 for begge analyserede år. Når M overstiger 50, vil de efficiente selskaber indgå i mere end 50 pct. af iterationer og dermed påvirke modellen i høj grad.

Uddybning af den specifikke analyse for egenskaben

Frontfastlæggelsen i Order-M bliver mindre påvirket af de efficiente selskaber og mere af de mindre effektive selskaber¹³ sammenlignet med en standard DEA model. Det medfører, at fronten i Order-M bliver mere fleksibel. En fleksibel front betyder, at fronten er mere kontinuer med færre hårde hjørner, og at for hvert hjørne dækker et mindre område, jf. Figur 3.5.

Figur 3.5 viser fronten for Order-M ved tre forskellige M'er samt en front i en standard DEA. Den sorte, standard DEA-front har fem frontselskaber og derfor også fem hjørner. Figuren viser, at Order-M-fronterne har færre hårde hjørner og derfor har en mere kontinuer udformning. Når M reduceres, er der større fleksibilitet i, hvilke selskaber der potentielt kan udgøre fronten, og dermed bliver Order-M-fronten mere afrundet.

Figur 3.5 Sammenligning af fleksibiliteten af fronten i DEA og Order-M



Anm.: Figuren viser frontfastsættelsen for henholdsvis DEA og Order-M ved tre forskellige værdier af M. De sorte prikker viser de enkelte selskaber.

Kilde: Egen tilvirkning

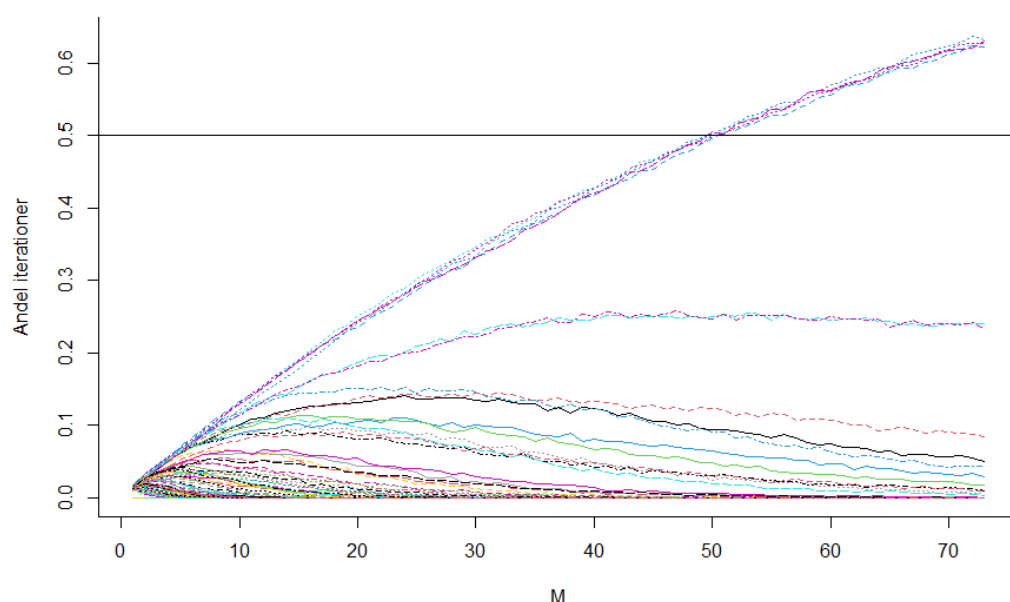
¹³ "Mindre effektive selskaber" betegner de vandselskaber som ikke har formået – i lige så høj grad som frontselskaberne – at omsætte omkostninger til produktion eller service.

Fleksibiliteten for Order-M betragtes som en fordel, da den tilnærmer sig formen på den forventede "sande", men ukendte front.

Udjævnningen af fronten medfører samtidig, at de efficiente selskaber påvirker resultaterne mindre, jf. Figur 3.6. Figuren viser andelen af iterationer, hvori hvert selskab indgår i frontfastlæggelsen i Order-M. Hver linje indikerer et specifikt selskab. Fx er der fem selskaber, som for høje værdier af M indgår på fronten i cirka 60 pct. af alle iterationerne.¹⁴ Det betyder, at de har stor indflydelse på den samlede Order-M-front. Når M reduceres, indgår disse selskaber i færre iterationer. Det skyldes, at sandsynligheden for at blive tilfældigt udtrukket falder.

Samtidig viser figuren, at andre selskaber indgår i flere iterationer, når M falder. Det skyldes, at andre selskaber kun kan ligge på fronten i de iterationer, hvor de efficiente selskaber ikke bliver trukket, samt at sandsynligheden for at trække et af de efficiente selskaber falder når M reduceres.

Figur 3.6 Andel af iterationer et selskab indgår på fronten



Anm.: Figuren viser andelen af Order-M-iterationer hvert selskab indgår i. Hver linje indikerer et selskab.

Kilde: Egen tilvirkning

Resultaterne fra Figur 3.6 kan bruges til at bestemme M. Forsyningssekretariatet vurderer, at fronten først bliver udjævnet tilstrækkeligt, når der ikke er nogle selskaber, der indgår i mere end 50 pct. af iterationerne (indikeret ved den horisontale linje). Det vil sige, at denne analyse

¹⁴ Bemærk, at de ville indgå i 100 pct. af iterationerne, hvis tilbagelægning ikke var tilladt.

viser, at M maksimalt skal være cirka 50. Bemærk, at det er tilfældigt, at 50 pct. af iterationer svarer til M lig med 50.¹⁵

4. Modellen opfylder betingelser og antagelser fra DEA

Sådan analyseres egenskaben

Egenskaben analyseres ved hjælp af to metoder:

1. For det første undersøges det, hvor mange selskaber der får en efficiensscore over 1 (herefter kaldet superefficient). Hvis der er for mange superefficiente selskaber, er det et tegn på, at Order-M ikke beregner retvisende efficiensscorer.
2. Hver iteration i Order-M skal opfylde litteraturens retningslinjer om, hvor mange observationer, der minimum skal bruges i en DEA-model.

Det viser resultaterne

Resultaterne viser, at antallet af superefficiente selskaber ligger forholdsvis stabilt for M større end 30. Derfor bør M på baggrund af den første delanalyse sættes til cirka minimum 30. Resultaterne for anden delanalyse viser, at der skal bruges minimum 13 selskaber i hver iteration for at efterleve litteraturens retningslinjer om antallet af observationer i en DEA-model.

Uddybning af den specifikke analyse for egenskaben

Order-M skal opfylde de samme egenskaber som DEA. Det betyder blandt andet, at selskaberne som udgangspunkt får beregnet en efficiensscore på maksimum 1. Herudover skal der være nok observationer i hver iteration fra Order-M til, at den underliggende DEA-model opfylder dens antagelser om produktionsmulighedsområdet.

i. Antal superefficiente selskaber

For det første er definitionen på en efficiensscore, at den maksimalt kan være 1. Hvis den er større end 1 (såkaldt superefficient), er et selskab mere effektivt, end hvad der bør være muligt¹⁶.

Der vil imidlertid altid være minimum ét selskab, der identificeres som superefficient i Order-M. Det skyldes, at frontselskaberne fra en standard-DEA altid vil blive vurderet som superefficiente i en Order-M-model, da Order-M kun kan give højere score end en standard DEA-model.

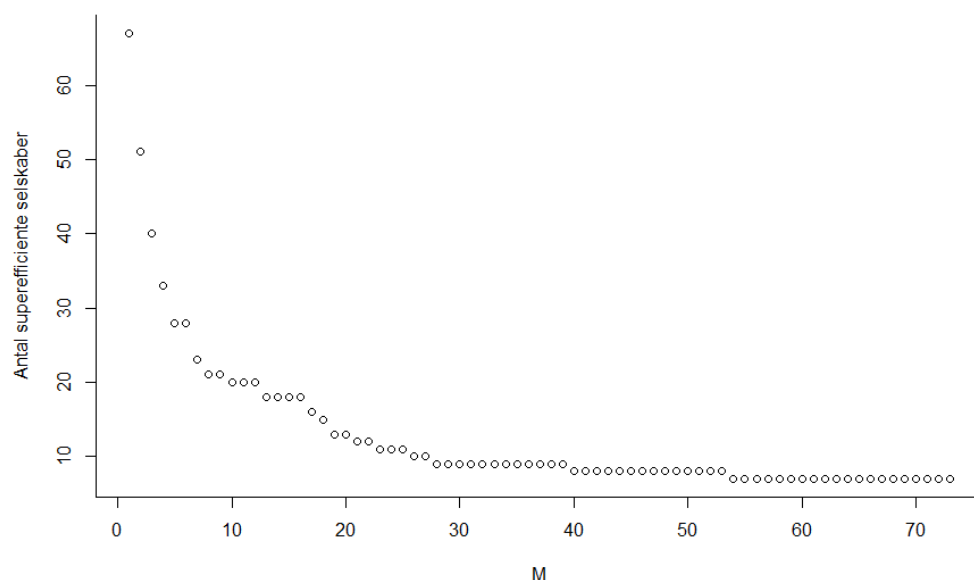
Det ønskes at minimere antallet af superefficiente selskaber. Forsyningssekretariatet vurderer ikke, at det er nødvendigt at vælge et M, der giver et absolut minimum antal superefficiente selskaber, men at antallet af superefficiente selskaber skal være relativt lavt.

Sammenhængen mellem antallet af superefficiente selskaber og M er vist i Figur 3.7. Figuren viser, at antallet af superefficiente selskaber ligger nogenlunde stabilt for M større end 28.

¹⁵ Sandsynligheden kan også beregnes med en lukket formel. Fx er sandsynligheden for at det efficiente selskab indgår på fronten $1 - \left(\frac{N-1}{N}\right)^M = 0,498 \text{ pct.}$ for $N = 73$ og $M = 50$, hvilket skyldes, at det efficiente selskab altid ligger på fronten, hvis det bliver udtrukket.

¹⁶ Bogetoft & Otto, 2010

Figur 3.7 Antal superefficiente selskaber



Anm.: Figuren viser antallet af superefficiente selskaber

Kilde: Egen tilvirkning

Forsyningssekretariatet vurderer på baggrund af første delanalyse for antallet af superefficiente selskaber, at M skal ligge på minimum 28.

ii. Minimumskrav for litteraturens retningslinjer

Der skal være nok observationer i hver iteration fra Order-M til, at den underliggende DEA-model opfylder de generelle antagelser om produktionsmulighedsområdet. I litteraturen er der en diskussion om, hvor mange observationer en DEA-model som minimum skal bruge, før man kan stole på resultaterne. Diskussionen spænder bredt, og der er ikke konsensus.

De tre mest anerkendte retningslinjer for antallet af observationer er, at antallet af observationer i en DEA-model skal være større end henholdsvis:

- » $2 \cdot I \cdot O$,
- » $2 \cdot (I + O)$, eller
- » $3 \cdot (I + O)$

- hvor I er antallet af inputs og O er antallet af outputs¹⁷.

¹⁷ Charles, Aparicio, & Zhu (2019)

I Forsyningssekretariatets benchmarkingmodel skal der derfor som et minimum bruges 9 observationer for at opfylde alle tre retningslinjer. Det skyldes, at der anvendes 1 input (totalomkostninger) og 2 outputs (OPEX- og CAPEX-netvolumenmål).

Hertil skal det bemærkes, at Forsyningssekretariatet har valgt at tillade, at det samme selskab kan trækkes flere gange ved udtræk af M antal selskaber til hver iteration i Order-M ("tilbagelæg"). Derfor kan der opstå dubletter, når der skal trækkes selskaber i en given iteration. Derfor bør M sættes til en værdi, der er større end 9. For at sikre en tilstrækkelig høj sandsynlighed for at trække minimum 9 unikke observationer i hver iteration skal M være minimum 13¹⁸.

På baggrund af anden delanalyse for minimumskravet til for litteraturens retningslinjer, skal M fastsættes til minimum 13 2023.

¹⁸ Med tilbagelægning er der en sandsynlighed for, at samme selskab trækkes flere gange, hvorfor M skal være større end 9 for med stor sandsynlighed at trække 9 unikke selskaber. Når M er lig 13, er der 99,9 pct. sandsynlighed for at have minimum 9 unikke selskaber i en given iteration.

Kapitel 4

Opsamling

I dette tekniske modelpapir er det vist, hvordan Order-M udvider DEA til at kunne tage hensyn til blandt andet støj, hvilket giver selskaberne en højere efficiensscore. Det medfører, at Order-M kan anvendes i stedet for at anvende en kombination af DEA og SFA.

Order-M er en simpel udvidelse til DEA. Order-M kræver, at der tages stilling til værdien af parameteren M – som er udtryk for, hvor mange selskaber, der skal indgå i hver iteration. Ved hjælp af fire opstillede egenskaber er der identificeret et spænd for M på omkring 30-50. Dette spænd er gældende for så vidt angår data fra benchmarkingen i 2023 og kan derfor ikke overføres direkte til en eventuel kommende benchmarkingmodel. Hvis Order-M implementeres, forventes det at bruge samme fremgangsmåde, og det forventes, at resultaterne fremover vil være nogenlunde ens ved anvendelse af fremtidige data. Forsyningssekretariatet har også foretaget analyserne på historisk data for drikkevand for 2021 samt spildevand for henholdsvis 2022 og 2024. Analyserne viser, at spændet for M er forholdsvis stabilt over årene. I en endelige benchmarkingmodel vil der skulle fastlægges et enkelt M inden for spændet, som vil skulle bruges til årets konkrete benchmarkingmodel.

Appendix A

Matematisk formulering af DEA og DEA-Order-M

DEA

Efficiensscorerne i DEA findes ved at løse følgende lineære optimeringsproblem:

$$\min_{\lambda, E} E^0 \quad (1)$$

$$E^0 x^0 \geq \sum_{k=1}^K \lambda^k x^k \quad (2)$$

$$y^0 \leq \sum_{k=1}^K \lambda^k y^k \quad (3)$$

Optimeringsproblemet løses for hvert selskab, hvor E^0 angiver selskabets efficiensscore, x^0 er selskabets input, K er antallet af selskaber i modellen, λ^k er en fri skaleringsvariabel, x^k er input for selskab k , y_0 er output for det givne selskab og y^k er output for selskab k . Bemærk, at inputs og outputs kan være vektorer, hvis der er mere end henholdsvis ét input og ét output. I Forsyningssekretariatet benchmarkingmodel er der fx to outputs, hvorfor ligning (3) gælder for både OPEX og CAPEX.

Ligning (1) angiver, at der først findes den laveste efficiensscore, der opfylder de to bi-betingelser. Ligning (2) angiver, at selskabets effektive omkostninger skal være lig med eller større end fronten. Det svarer til, at selskab D i Figur 2.1 følger den stiplede linje indtil den rammer fronten. Ligning (3) angiver, at selskabets output er mindre end eller lig med outputtet for det punkt på fronten, som det sammenlignes med.

Ovenstående optimeringsproblem er baseret på en inputorienteret DEA-model med konstant skalaafkast, hvilket er det, Forsyningssekretariatet bruger i benchmarkingmodellen i vandreguleringen.¹⁹

¹⁹ For andre valg af orienteringer og antagelser om skalaafkast se fx (Bogetoft & Otto, 2010)

DEA-Order-M

Efficiensscoren i Order-M findes ved at tage gennemsnittet af N DEA-modeller, hvor hver DEA-model køres på en delmængde af alle selskaberne.

I hver iteration løses følgende lineære optimeringsproblem:

$$\min_{\lambda, E} E_i^0 \quad (4)$$

$$E_i^0 x^0 \geq \sum_{k \in \bar{M}_i} \lambda^k x^k \quad (5)$$

$$y^0 \leq \sum_{k=1} \lambda^k y^k \quad (6)$$

Optimeringsproblemet løses for hvert selskab, hvor parametrene er de samme som i DEA, jf. ligning (1) - (3). Den eneste forskel er, at antallet af selskaber (K) i DEA nu er byttet ud med den udtrukne gruppe af selskaber, $\bar{M}_i$. Gruppen $\bar{M}_i$ findes ved at trække M tilfældige selskaber med tilbagelægning fra hele populationen af selskaber. Bemærk, at efficiensscoren nu er indekseret med i, som refererer til den givne iteration.

Ovenstående optimeringsproblem løses for hvert selskab N gange, hvor $\bar{M}_i$ genberegnes for hver iteration. Den endelige Order-M efficiensscore for selskab 0 er givet ved gennemsnittet af efficiensscorerne for alle iterationer, jf. ligning (7).

$$E^0 = \frac{\sum_{i=1}^N E_i^0}{N} \quad (7)$$

Referencer

- Bogetoft, P., & Otto, L. (2010). *Benchmarking with DEA, SFA, and R*. Springer Science & Business Media.
- Cazals, C., Florens, J.-P., & Simar, L. (2002). Nonparametric frontier estimation: a robust approach. *Journal of Econometrics*, 106(1), s. 1-25.
- Charles, V., Aparicio, J., & Zhu, J. (2019). The curse of dimensionality of decision-making units: A simple approach to increase the discriminatory power of data envelopment analysis. *European Journal of Operational Research*, s. 929-940.
- Charnes, A., Cooper, W. W., & Rhodes, E. (1978). Measuring the efficiency of decision making units. *European Journal of Operational Research*, 2(6), s. 429-444.
- Daraio, C., & Simar, L. (2007). Conditional nonparametric frontier models for convex and nonconvex technologies: a unifying approach. *Journal of productivity analysis*, 28(1), s. 13-32.
- Mann, H. B., & Whitney, D. R. (1947). On a Test of Whether one of Two Random Variables is Stochastically Larger than the Other. *The Annals of Mathematical Statistics*, 18(1), s. 50-60.
- Silva, T. D., Martins-Filho, C., & Ribeiro, E. (2016). A comparison of nonparametric efficiency estimators: DEA, FDC, DEAC, FDCH, order-m and quantile. *Economic Bulletin*, 36(1).
-