

# Praktiske erfaringer fra estimering med usikkerhed i IT projekter

Estimering af IT projekter har gennem tiderne altid været en særdeles vanskelig disciplin, og der findes næppe den eller de metoder der er den gyldne løsning på problemet.

I følge en hollandsk undersøgelse overskred 80% af alle IT projekter budgettet og/eller kalendertiden med mere end 50% [1]. Tallene taler for sig selv, og stemmer vel godt overens med manges praktiske erfaringer.

Denne artikel beskriver kort principperne bag estimering med usikkerhed ud fra en pragmatisk indfaldsvinkel og viser herefter eksempler på de opnåede resultater i praksis. Principperne her i artiklen er baseret på praktiske erfaringer og empiriske data.

Netop estimering med usikkerhed kan være med til at forbedre et IT projekts estimat væsentlig og dermed øge chancen for, at projektet opleves som en succes.

## Hvad er et estimat ?

Hvis man beder en projektleder om at estimere et IT projekt enten i timer og/eller kroner vil de fleste vende tilbage med svaret, at projektets budgetterede indsats er på 6300 timer og den budgetterede omkostning er på 5,8 millioner kroner.

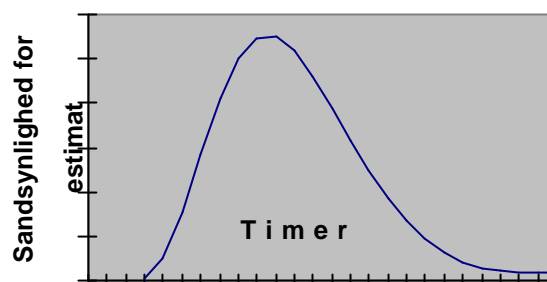
*Er det et godt svar på spørgsmålet ?*

Det er det næppe, da man intuitivt kan sige sig, at estimater må være forbundet med en større eller mindre usikkerhed afhængig af hvor godt opgaven er beskrevet, om der er tale om ny eller kendt teknologi osv.

Det er derfor rimeligt at antage, at et estimat er en stokastisk variabel [2], hvilket betyder:

*“Tilfældig, men til at forudberegne statistisk med tilnærmelsesvis sikkerhed”.*

Det viser sig endvidere at estimering af aktiviteter ofte følger en betafordeling, som vist på figur 1 [3,4]:



Figur 1 - Betafordeling

Betafordelingen er en såkaldt skæv fordeling, og den viser, at det altid vil tage en vis minimum tid at udføre en aktivitet (meget lille sandsynlighed), mens det kan tage næsten uendelig lang tid at udføre en aktivitet – I praksis må der dog være en endelig tid !

## Anvendelse af statistiske variable ved estimering

*Men hvordan bruger vi den viden i praksis, at estimering af aktiviteter ofte følger en betafordeling ?*

Både successiv kalkulation [3], og PERT<sup>1</sup> teknikken [4] arbejder med at angive 3 estimater nemlig:

- Optimistisk estimat ( $V_0$ )
- Mest sandsynligt estimat ( $V_S$ )
- Pessimistisk estimat ( $V_P$ )

For begge metoder kan man vælge at tage udgangspunkt i en betafordeling.

<sup>1</sup> PERT betyder Program Evaluation and Review Technique

Med baggrund i disse 3 estimater beregnes middelværdi ( $\mu$ ) og standardafvigelse ( $\sigma$ ) ud fra tilnærmede formler:

$$\mu = (V_0 + 3V_S + V_P) / 5$$

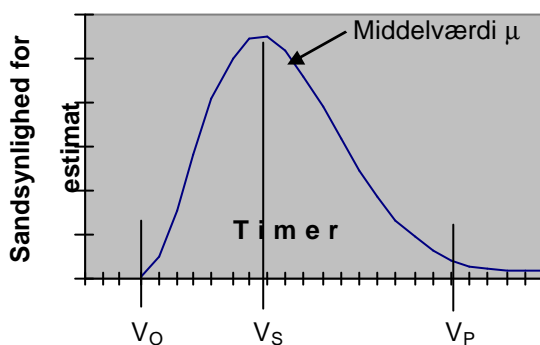
$$\sigma = (V_P - V_0) / 5$$

Følgende eksempel viser anvendelsen af ovennævnte formler:

$V_0$	$V_S$	$V_P$	$m$	$s$
15 [t]	25 [t]	40 [t]	26 [t]	5 [t]

Tabel 1 – Beregning af  $m$  og  $s$

Værdierne i tabel 1 er vist grafisk på figur 2 (dog ikke  $\sigma$ ).



Figur 2 - Betafordeling med værdier

Afhængig af det ønskede konfidensinterval kan usikkerheden beregnes

Konfidens-Interval	Formel (Interval)	Beregning Af Usikkerhed
68% (16%-84%)	$(\mu - \sigma) - (\mu + \sigma)$	21 – 31 [t]
95% (2,5%-97,5%)	$(\mu - 2\sigma) - (\mu + 2\sigma)$	16 – 36 [t]
99% (0,5% - 99,5%)	$(\mu - 3\sigma) - (\mu + 3\sigma)$	11 – 41 [t]

Tabel 2 - Beregning af usikkerhed

Tabellen skal forstås således, at hvis man vil være 95% sikker på at estimaterne holder skal man vælge intervallet 16 – 36 timer.

Man skal dog være varsom med for sikre antagelser på basis af tabel 2 på grund af:

- Der er tale om tilnærmede formler.
- Hvor valide er de 3 grundtal ( $V_0$ ,  $V_S$  og  $V_P$ ) ?

## Praktisk estimering med usikkerhed

Kan de viste beregninger forenkles ?

Selvom de tidligere viste beregninger er forholdsvis enkle er det min erfaring, at de kan forenkles yderligere.

Det er især baseret på den kendsgerning, at der i praksis er en stor skævhed i betafordelingen således at sandsynligheden for den optimistiske værdi er meget lille.

Derfor vælger jeg at forenkles estimaterne til 2 punkter på kurven kaldet estimat 1 ( $E_1$ ) og estimat 2 ( $E_2$ ).

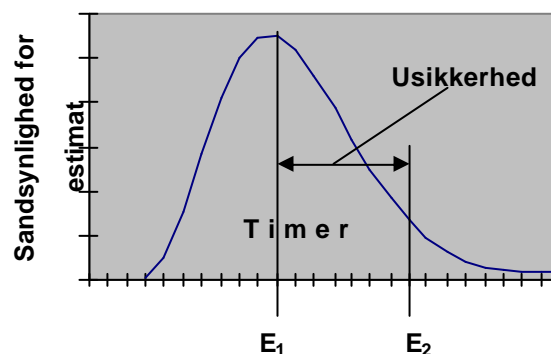
De 2 punkter kan eksempelvis findes ved følgende spørgsmål:

1. Hvor lang tid tager det typisk at udføre en given aktivitet ? (*dette benævnes estimat 1*)
2. Hvis tingene går galt hvor stor en usikkerhed<sup>2</sup> i procent skal der lægges til ? (*dette benævnes usikkerhed og forkortes U*)

Estimat 2 beregnes således:

$$E_2 = E_1 (1 + U[\%])$$

Figur 3 viser begreberne grafisk:



Figur 3– Estimat 1, Usikkerhed og Estimat 2 vist grafisk på betafordelingen

<sup>2</sup> Bemærk, at usikkerhed her mere bruges intuitivt og ikke i relation til konfidensintervaller som vist tidligere.

Estimat 1 og Estimat 2 kaldes bevidst ikke for henholdsvis mest sandsynlig og pessimistisk estimat, da man reelt ikke ved, hvor på betafordelingen de 2 punkter befinder sig.

Endvidere opfattes en pessimistisk værdi ofte som en slags "worst case" værdi, hvilket desværre ikke er tilfældet, da det i praksis kan gå meget værre end det der kaldes "worst case" (betafordelingen går som vist i figur 3 mod uendelig).

Det har vist sig at være meget hensigtsmæssigt at operere med estimat 1 og estimat 2, og så opgive en opgaves varighed til et interval (det vil sige fra  $E_1$  til  $E_2$ ).

Netop det at opgive et interval i stedet for et enkelt tal er en god praktisk måde at kommunikere et estimat til en styregruppe for et projekt, da de herved bliver klar over, at det ikke er realistisk at opgive et eksakt tal.

Metoden kan anvendes på både indsats i såkaldte effektive<sup>3</sup> timer, men også til estimering af omkostninger i IT projekter (eksempelvis maskinel, licenser, rejseomkostninger med flere).

### **Et simplificeret eksempel hvor metoden anvendes**

Arbejdsgangen for at etablere et estimat med usikkerhed er således:

1. Fastlæg omfanget af opgaven der skal estimeres og beskriv forudsætningerne for opgaven.
2. Nedbryd opgaven i aktiviteter så  $E_1 \leq 40$  timer for de enkelte aktiviteter<sup>4</sup>. Der er tale om en tommelfingerregel, som har vist sig at være nyttig i praksis. Denne nedbrydning kaldes på engelsk "work breakdown structure" – WBS.
3. Estimat 1 og usikkerheden fastlægges nu for opgaverne på nederste niveau (trin 2 og trin 3 gennemføres iterativt).
4. Der udledes et total estimat med usikkerhed ved simpel aritmetisk summering.

<sup>3</sup> Med effektive timer menes, at det er den tid en medarbejder ville anvende, hvis medarbejderen arbejdede 100% på opgaven, hvilket vil sige eksklusiv egen administration, gruppemøder m.v.

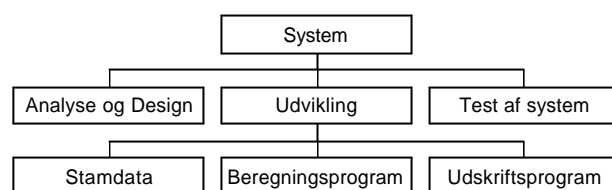
<sup>4</sup> Der kan være situationer, hvor det ikke er muligt at nedbryde opgaven så meget. I disse situationer vil man ofte kompensere den manglende nedbrydning med en større usikkerhed.

### **Trin 1 og trin 2**

Der skal udvikles et system til beregning af årlige omkostninger for personbiler, og det antages at bestå af følgende aktiviteter<sup>5</sup>:

- Analyse og design af systemet
- Udvikling af stamdata program
- Udvikling af beregningsalgoritme
- Udvikling af udskrivningsprogram
- Test af systemet.

Aktiviteterne hænger grafisk sammen som vist på figur 4 (work breakdown structure):



**Figur 4 - Work breakdown structure**

### **Trin 3 og trin 4**

I den følgende tabel 3 er estimat 1, usikkerhed og estimat 2 angivet på nederste niveau i WBS strukturen:

Aktivitet	$E_1$ [t]	U [%]	$E_2$ [t]
Analyse og Design af system	30	50%	45
Udvikling af stamdata	12	25%	15
Udvikling af beregningsalgoritme	36	75%	63
Udvikling af udskriftsprogram	18	25%	23
Test af systemet	12	50%	18
<b>Total</b>	<b>108</b>	<b>51%</b>	<b>164</b>

**Tabel 3 - Estimat med usikkerhed**

Af tabel 3 fremgår, at estimatet for udvikling af systemet er i intervallet fra 108 til 164 timer.

Totalen for Estimat 2 ( $E_2$ ) er fundet ved aritmetisk summering af de enkelte  $E_2$ . Det kan her diskuteres hvorvidt det er statistisk korrekt, da det afhænger af hvordan de enkelte aktiviteter er relateret til hinanden (korrelations faktoren).

<sup>5</sup> Der er absolut tale om et tænkt eksempel som ikke kan bruges til virkelige estimater for sådan en opgave.

Jeg har valgt at antage at samtlige aktiviteter er 100% korreleret, da det er min erfaring fra praksis, at hvis noget går galt har det en tendens til at påvirke mange aktiviteter i samme negative retning.

Det ville blive en noget teoretisk situation at skulle beskrive korrelations faktoren mellem de forskellige aktiviteter og så beregne totalen for Estimat 2 baseret på dette.

Som det ses af ovennævnte er der gjort meget for at forenkle den praktiske metode til estimering med usikkerhed, hvilket betyder:

- At metoden er forholdsvis nem at forstå og den kræver ikke et detaljeret kendskab til statistiske begreber.
- At den er meget simpel at implementere i regneark.

Begge dele gør, at metoden nemt kan bruges i praksis.

Som eksempel fra en større dansk virksomhed kan nævnes, at en fornuftig skabelon<sup>6</sup> i et regneark kan udvikles og afprøves på ca. 100 timer, og projektledere kan uddannes i brugen af regnearket på cirka 2 timer.

### **Sammenhæng til andre estimeringsmetoder**

Metoden til estimering med usikkerhed (herefter forkortet EMU) kan nemt kombineres med en række andre estimeringsmetoder, og kan betragtes som et vigtigt supplement til disse.

Der findes en række metoder til estimering af IT projekter [5], og her følger en kort beskrivelse af disse samt sammenhængen til EMU.

#### **Top-down & Bottom-up**

Uanset om der vælges en top-down eller bottom-up metode til fastlæggelse af WBS strukturen vil den naturligvis kunne kombineres med EMU.

Det vil typisk være således, at størrelsen af usikkerheden er mindre ved bottom-up end ved top-down metoden.

---

<sup>6</sup> Skabelonen indeholder betydelig mere end blot det der fremgår af tabel 3. Eksempelvis summeringer pr. fase, checklister, omregning til omkostninger ændringsstyring med mere.

Grunden til at usikkerheden normalt er mindre ved bottom-up metoden er, at en detaljeret nedbrydning af en opgave gør, at vi nemmere kan overskue og estimere opgaven, og dermed reducere den reelle usikkerhed.

#### **Ekspert metoden**

Fastlæggelse af estimat 1 og estimat 2 i eksemplet fra tidligere er fundet ud fra estimatorens erfaringer.

Dette kan kaldes ekspert vurderinger eller blot estimering baseret på erfaringer.

Langt de fleste estimater baseres i større eller mindre grad på erfaringer, så det er nok den mest almindelige estimeringsmetode.

Denne metode hænger naturlig sammen med EMU, som vist i eksemplet.

#### **Analogi metoden**

Analogi metoden baserer sig på at ekstrapolere estimater for et nyt projekt ud fra estimater fra tidligere gennemførte projekter korrigeret for forskellige forhold som kompleksitet, tekniske forhold med flere.

Analogi metoden kan bruges direkte til at finde estimat 1, mens usikkerheden kan udtrykke de variationer (usikkerheder) der skønnes at være til stede.

#### **Algoritmiske modeller**

Algoritmiske modeller kaldes også estimering af omkostninger ved hjælp af parametre.

Fælles for disse metoder er, at de baserer sig på en parametrisk sammenhæng mellem størrelsen af et system (S) og estimatet (E) i timer og/eller kroner, som vist nedenfor:

$$E = \text{funktion}(S)$$

Denne sammenhæng mellem E og S kan være lineær, eksponentiel med flere og indeholde en lang række konstanter [6].

Estimering ved hjælp af function point [7] og source kodelinier (eksempelvis COCOMO [5]) er eksempler på algoritmiske modeller.

Det er oplagt at bruge historiske data fra eksemplvis Function Point metoden sammen med EMU.

Lad os antage, at en IT organisation har fundet ud af, at det koster gennemsnitlig fra 14 – 22 timer at gennemføre 1 Function Point (FP).

Vi har et program der indeholder 17 FP, så kunne Estimat 1 og Estimat 2 beregnes således:

$$E_1 = 14 * 17 = 238 \text{ timer}$$

$$E_2 = 22 * 17 = 374 \text{ timer}$$

Her vil det være korrekt og nødvendigt at se bort fra tommelfingerreglen om, at  $E_1 \leq 40$ , da en yderligere nedbrydning ikke er relevant.

### **Faldgruber ved estimering med usikkerhed**

"Der er ingen roser uden torne", hvilket også gælder for Estimering med usikkerhed (EMU).

Mange af de faldgruber der er ved EMU gælder generelt ved estimering af IT projekter. De er dog medtaget så estimeringen kan sættes i det rette perspektiv.

Eksempler på faldgruber er:

- Den sunde fornuft sættes til side.
- Skævhed i fastlæggelse af estimat 1 og estimat 2.
- Omfanget af opgaven ændrer sig.
- Usikkerhed og risikostyring "blandes" sammen.

#### **Den sunde fornuft sættes til side**

Der er en tendens til, at vi bliver fanget af tallenes magt, når først vi har etableret metoder og værktøjer (som her til estimering).

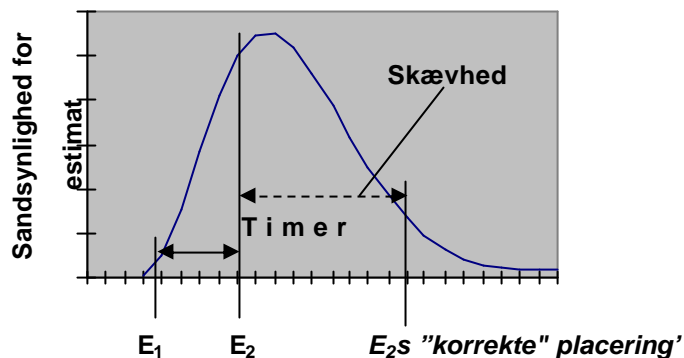
Det betyder, at vores almindelig dømmekraft erstattes af tal og formler.

En god regel er at være flere om at estimere en opgave, og aldrig stole på tallene alene – brug din sunde fornuft som en slags "sanity" check.

### **Skævhed i fastlæggelse af estimat 1 og estimat 2**

Det er en kendt sag at næsten alle er for optimistiske når de estimerer. Jeg plejer at kalde det optimistfaktoren, og den er en stor risiko i projekter, og giver sig udslag i en skævhed (*engelsk bias*).

Skævheden viser sig ved at estimat 1 og estimat 2 er forskudt i betafordelingen som vist på figur 5.



Figur 5 – Optimistfaktoren vist som skævhed på Betafordelingen

Det fremgår umiddelbart af figur 5, at denne skævhed kan give store problemer ved fastlæggelsen af slutdatoen for projektet og projektets totale omkostninger.

#### **Omfanget af opgaven ændrer sig**

Trin 1 i arbejdsgangen for estimering beskæftiger sig med fastlæggelse af omfanget af opgaven.

Hvis de forudsætninger der måtte være stillet op for omfanget af opgaven brister<sup>7</sup> må disse ændringer afspejles i estimatet (reestimering).

EMU metoden tager ikke direkte højde for bristede forudsætninger, tilføjelse af mere funktionalitet i et IT system osv.

Derimod tager EMU metoden højde for de naturlige variationer der er for at gennemføre en given aktivitet.

<sup>7</sup> Hvilket den typisk gør i alle IT projekter. Dette skal håndteres med en effektiv ændringsstyring.

Der er flere måder der kan tage højde for ovennævnte (og de kan bruges i flæng):

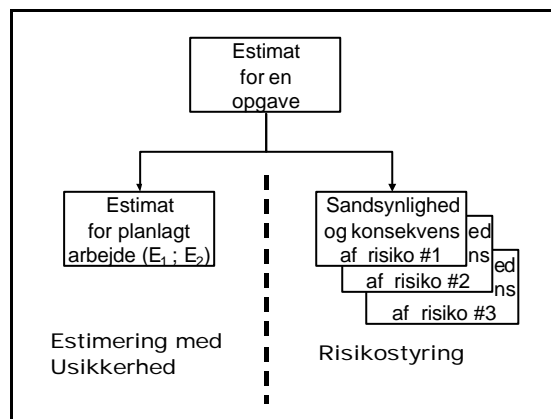
- Den første er meget klart at beskrive hvilke forudsætninger der er for opgaven, så det ikke er manglende klarhed der giver problemet.
- At indbygge en reserve til uforudsete opgaver [8].
- At gennemføre ændringsstyring på projektet så konsekvenser af ændringer hele tiden fastlægges og der tages aktiv stilling til om ændringer skal gennemføres.
- At gennemføre en kvantitativ risikostyring som indarbejder "risikoen" for, at opgavens omfang ændres.

### Usikkerhed og risikostyring "blandes" sammen

Der kan være en god grund til at skelne mellem begreberne usikkerhed og risikostyring, så her følger en forklaring på forskellen:

- Usikkerhed skal opfattes som den naturlige variation der er for en opgaves udførelse (som nævnt tidligere følger det typisk en betafordeling – se eksempelvis figur 1).
- Risikostyring skal opfattes som håndtering af en række risikofaktorer der kan indtræffe eller udeblive og dermed ændre et estimat drastisk (måske mange hundrede procent).

De 2 begreber er vist på figur 6:



Figur 6 – Illustration af usikkerhed og risikostyring

En uddybende behandling af risikostyring er udenfor denne artikels rammer se eksempelvis reference 9.

## Erfaringer fra anvendelsen af metoden i konkrete IT Projekter

Estimering med usikkerhed (EMU) som beskrevet i denne artikel anvendes i flere danske virksomheder og her følger erfaringer fra 4 projekter.

De 4 projekter fremstår anonyme, og der anvendes fortrinsvis relative talstørrelser i stedet for absolutte tal for at bibeholde anonymiteten.

Projekternes størrelse udtrykt i kroner varierer fra 3 til 30 millioner kroner.

Det der måles på i de konkrete projekter er 2 parametre:

1. Hvad er forskellen på den estimerede indsats i timer og den realiserede indsats i timer ? (Dette kaldes  $CV_{TIMER}$  for Cost Variance i Earned Value Metoden [8]).
2. Hvad er forskellen på den estimerede omkostning og den realiserede omkostning i kroner ? Dette kaldes  $CV_{KRONER}$  for Cost Variance i Earned Value Metoden [8]).

Beskrivelsen er herefter opdelt i 2 afsnit:

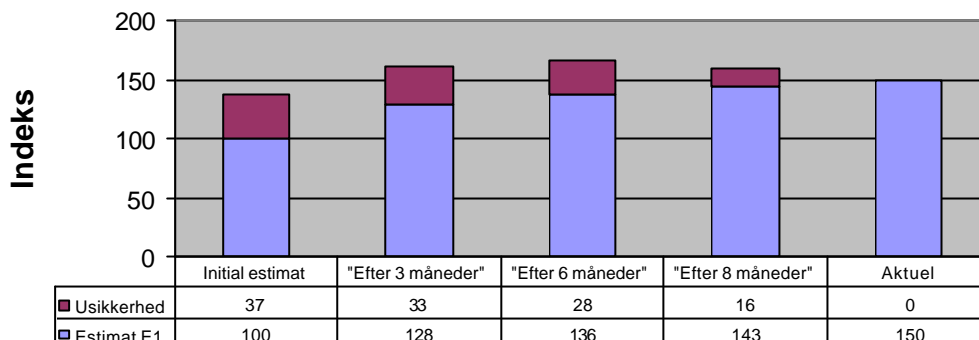
- Et afsnit som beskriver et enkelt projekts forløb indenfor de 2 parametre
- Et afsnit som beskriver nøgletallene for de 4 projekter

### Et projekt fra virkeligheden, hvor EMU metoden er anvendt.

Figur 7 på næste side viser henholdsvis hvordan estimat 1 og usikkerhed har ændret sig over en periode.

Figuren skal forstås således, at de enkelte stave i histogrammet viser resultatet af de reestimeringer der er udført (her er kun medtaget udvalgte reestimeringer. I virkeligheden er dette projekt blevet reestimeret cirka 1 gang om måneden).

### Estimat i timer



Figur 7 – Estimat i timer vist over en periode

Af figur 7 ses, at intervallet (dvs. fra E<sub>1</sub> til E<sub>2</sub>) varierer en del fra initial estimat (100 – 137 enheder) til sidste estimat her kaldet "Efter 8 måneder" (143 – 159 enheder).

Det der er sket er, at der er bygget mere og mere erfaring ind i estimaterne i takt med at projektet er blevet gennemført.

Beregning af variansen i timer er således:

$$CV_{\text{TIMER}} = (E_1 + U) - \text{Aktuel}$$

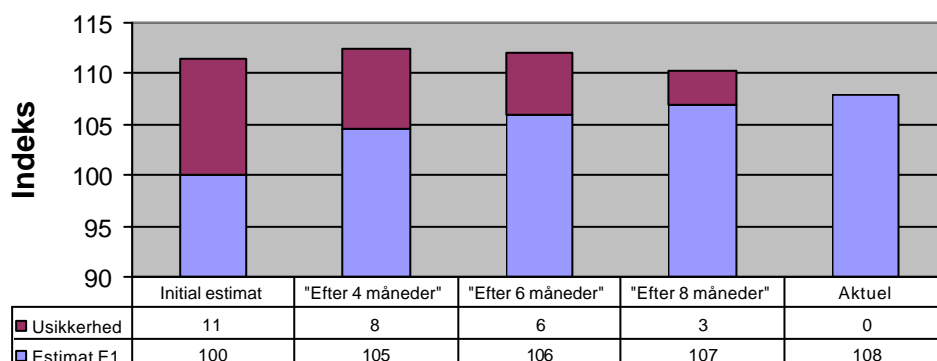
$$CV_{\text{TIMER}} = (100 + 37) - 150$$

$$CV_{\text{TIMER}} = - 13 \text{ [Enheder]}$$

$$CV_{\text{TIMER}} = - 9,5 \%$$

Det betyder, at projektet er blevet estimeret 10% for lidt i forbindelse med initial estimatet inklusive usikkerhed.

### Estimat i Kroner



Figur 8 - Estimat i kroner vist over en periode

På figur 8 vises de tilsvarende tal for omkostninger udtrykt i kroner.

Figur 7 og figur 8 kan ikke direkte sammenlignes da opgørelsesmetoden for timer og kroner er forskellig i det pågældende projekt.

EMU metoden er dog anvendt i både estimeringen af timer og kroner, så set fra en metode synsvinkel er det underordnet.

Ved beregning af variansen i kroner når vi frem til følgende resultat:

$$\begin{aligned} CV_{\text{KRONER}} &= (E_1 + U) - \text{Aktuel} \\ CV_{\text{KRONER}} &= (100 + 11) - 108 \\ CV_{\text{KRONER}} &= 3 \text{ [Enheder]} \\ CV_{\text{KRONER}} &= 2,7 \% \end{aligned}$$

Her er der tale om at det aktuelle forbrugte har været en anelse mindre end det estimerede.

Det er et resultat som de fleste vil betragte som udmærket, da historien er fyldt med eksempler på projekter med store budget overskridelser [1].

### **Nøgletal for de 4 projekter**

Her følger en tabel over nøgletallene for de 4 projekter:

	CV <sub>TIMER</sub>	CV <sub>KRONER</sub>
Projekt 1	-9,5%	2,7%
Projekt 2	<i>Mangler</i>	36,4%
Projekt 3	11,6%	27,9%
Projekt 4	17,3%	<i>Mangler</i>

**Tabel 3 - Nøgletal for 4 projekter**

For alle projekter gælder følgende:

- Sammenligningen er foretaget mellem initial budget inklusive usikkerhed (E2) og de aktuelle tal.
- Eventuelle projektændringer er holdt ude af tallene, så der udelukkende måles i forhold til initial omfang.

Af tabel 3 ses, at flere projekter holder sig både under budgettet i timer og kroner.

## **Konklusion**

EMU metoden er et praktisk redskab til at forbedre estimeringsprocessen i IT projekter<sup>8</sup>.

Praktiske erfaringer viser, at estimering af IT projekter kan forbedres væsentligt ved at anvende EMU metoden.

EMU metoden er samtidig forholdsvis enkel og kan håndteres i et almindelig regneark, så her er en praktisk metode der nemt kan tages i anvendelse.

### **Referencer:**

1. Kostenbeheersing bij automatiseringsprojecten: een empirisch onderzoek, Informatie, Vol. 31, no. 1. - F.J. Heemstra, W. Siskens and H. van der Stelt (1989)
2. Controlling Software Projects – Tom De Marco - ISBN 0-13-171711-1 (1982)
3. Rapport over successiv kalkulation – Steen Lichtenberg (1971)
4. A guide to Project Management Body of Knowledge - PMI Standards Committee - ISBN 1-880410-12-5 (1996)
5. Software Engineering Economics – Barry Boehm - ISBN 0-13-822122-7 (1981)
6. Parametric Cost Estimating Handbook – Joint Government / Industry Initiative – Fall 1995 – Tilgængelig på Internet [www.contracts.hq.navsea.navy.mil/webdata/pceh](http://www.contracts.hq.navsea.navy.mil/webdata/pceh)
7. Applied Software Measurement – Capers Jones - ISBN 0-07-032826-9 (1996)
8. Earned Value Project Management - Quentin W. Fleming & Joel M. Koppelman - ISBN 1-880410-38-9
9. Practical Risk Assessment for Project Management – Stephen Grey - ISBN 0-471-93979 (1995)
10. Proactive Management of Uncertainty Using the Successive Principle – Steen Lichtenberg ISBN 87-502-0822-5

<sup>8</sup> Fokus i denne artikel har været på IT Projekter, men EMU metoden er også anvendelig indenfor andre projektyper.