

## FORORD

Denne hovedfagsoppgaven er en del av Cand.scient graden ved Institutt for informatikk, studieretning for systemarbeid ved Universitetet i Oslo. Arbeidet ble påbegynt så tidlig som våren 2000, og har således gjennomgått en lang og god modningsprosess.

Jeg vil takke min arbeidsgiver Høgskolen i Gjøvik, for å ha lagt til rette for å gjennomføre studiene ved siden av min stilling som høgskolelærer. Dette har de gjort ved blant annet å tildele et høyskolelektor-kvalifiseringsstipend. Avdelingsledelsen ved avdeling for teknologi har vært fleksible med å gi undervisningsfrie perioder, slik har vært mulig å fullføre arbeidet med oppgaven.

Magne Jørgensen har vært min veileder. Han har gjennom hele prosessen vært en meget inspirerende diskusjonspartner og fortjener takk for dette.

Jeg har gjennom de siste tre og et halvt år gjort utallige henvendelser til biblioteket ved Høgskolen i Gjøvik, med forespørsler om å skaffe fagartikler og lærebøker i tilknytning til hovedfagsarbeidet. Jeg takker for at de alltid har gjort sitt ytterste for skaffe materialet.

Mine tre sønner og min kone måtte nøye seg med en svært begrenset ”lærerferie” fra min side denne sommeren. Takk for at dere tok dette med godt humør og en god porsjon omsorg for ”han far som stadig måtte til Gjøvik en tur”.

Gjøvik, juli 2003.

Tom Røise

## SAMMENDRAG

Estimater er viktig styringsinformasjon i systemutviklingsprosjekter. Prosjektledere og virksomhetsledelsen ønsker en tidlig og treffsikker prediksjon på omfanget av ressurser som trengs til å utvikle et datasystem. I forsøket på å imøtekomme dette har det vært forsket på og lansert en rekke alternative estimeringsmetoder innen fagfeltet.

Empiriske funn viser at det innen systemutvikling ikke finnes noen estimeringsmetode som er best uavhengig av kontekst. Det viser seg at de alternative estimeringsmetodene har styrker og svakheter som slår forskjellig ut når metodene anvendes i ulike organisasjons- og prosjektkontekster. Det er interessant å se nærmere på hvordan man kan utnytte dette til å gjøre et bevisst valg av estimeringsmetode ved starten av et prosjekt. Det finnes svært lite forskning som spesifikt har sett på valg av estimeringsmetode.

I dette hovedfagsarbeidet har jeg utarbeidet et rammeverk som skal hjelpe prosjektledere og andre estimeringsansvarlige til å velge den mest egnede estimeringsmetoden. I dette rammeverket står ivaretagelse av prosjekt- og organisasjonskontekst svært sentralt. Arbeidet er basert på litteraturstudier av publisert forskning innen estimering av systemutvikling supplert med publikasjoner på prediksjonsteori fra andre fagområder.

Rammeverket er basert på valgstrategien *Strukturert Vurdering*. Her vurderes de eksisterende estimeringsmetodene ut fra et sett utvalgte valgkriterier. Rammeverket gir dermed et innblikk i hva som finnes av alternative metoder. Dette er gunstig da kunnskapen om dette er lav blant mange prosjektledere og estimeringsansvarlige. Litteraturstudiet viste at de eksisterende metodekategoriseringene dekker modellbaserte metoder godt, men at bredden innen ekspertvurderinger ikke er ivaretatt. Jeg har derfor lansert en ny kategorisering med ti sub-kategorier av estimeringsmetoder. Disse dekker så vel modellbaserte metoder som ekspertvurderinger på en mer balansert måte.

Treffsikkerhet har hittil vært det dominerende kriteriet for å bedømme estimeringsmetoder. For å ivareta konteksten på en god måte, foreslår jeg i rammeverket at tretten ulike valgkriterier vurderes når man skal velge blant de ti aktuelle estimeringsmetodene. Kriteriene skal sikre at man tar hensyn til:

- Begrensinger og rammer for estimeringsprosessen
- Formålet med og anvendelsen av estimatet
- Evne og kapasitet til å estimere
- Prosjektkarakteristika
- Tilgjengelighet på informasjon

Ved å benytte kriteriene vil konteksten i stor grad sette sitt preg på metodevalget. Det er lagt vekt på at rammeverket skal være tilpasset det faktum at estimering er en støtteaktivitet innen systemutvikling med en begrenset tilgang på estimeringsressurser.

# INNHold

<b>1. INNLEDNING .....</b>	<b>4</b>
1.1. Tema .....	4
1.2. Problemstilling.....	5
1.3. Avgrensning.....	6
1.4. Mål og målgruppe .....	6
1.5. Metodevalg .....	7
1.6. Oppgavens struktur .....	9
<b>2. SENTRALE TREKK FRA LITTERATURSTUDIEN .....</b>	<b>11</b>
2.1. Estimering – en støtteaktivitet ved utvikling av datasystemer .....	11
2.2. Begrepsbruk .....	13
2.3. Valg av estimeringsmetode – et lite omtalt tema i litteraturen .....	13
2.4. Metodevalgets plassering i estimeringsprosessen.....	18
2.5. Metodenes forening av analyse og intuisjon.....	22
2.6. Kombinasjon av metoder .....	24
<b>3. STRATEGI FOR VALG AV ESTIMERINGSMETODE .....</b>	<b>25</b>
3.1. Tilnærminger til å velge metode .....	25
3.2. Gjennomgang av seks strategier for metodevalg.....	25
3.3. Bruk av <i>Strukturert Vurdering</i> i rammeverket .....	28
<b>4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER .....</b>	<b>30</b>
4.1. Hovedkategorier.....	31
4.2. Status for hver av de tre hovedkategoriene.....	32
4.3. Subkategorisering av metodene .....	36
4.4. Modellbaserte metoder .....	36
4.5. Analogibaserte metoder .....	41
4.6. Ekspertvurderinger.....	42
<b>5. KRITERIER FOR METODEVALG.....</b>	<b>53</b>
5.1. Kriterier som tidligere er omtalt i publisert litteratur .....	54
5.2. Forslag til valgkriterier .....	56
5.3. Detaljertbeskrivelse av kriteriene .....	58
5.4. Status for rammeverket.....	74
<b>6. ANVENDELSE AV RAMMEVERKET .....</b>	<b>75</b>
6.1. Utkast til anvendelsesveiledning .....	75
6.2. Valg av metode på organisasjonsnivå og prosjektnivå.....	79
<b>7. KONKLUSJON OG FORSLAG TIL VIDERE ARBEID .....</b>	<b>80</b>
7.1. Konklusjon.....	80
7.2. Forslag til videre arbeid .....	81

# 1. INNLEDNING

## 1.1. Tema

Estimering av kostnader og innsatsbehov utgjør en del av den totale prosjektstyringen i systemutviklingsprosjekter. Estimering dreier seg her om å gi gode prediksjoner på omfanget av de ressurser som trengs til å ferdigstille en programvare. Blant forskere så vel som blant prosjektledere og andre som utøver estimering i praksis, er det enighet om at gode estimater utgjør en meget nyttig styringsinformasjon. Desto tidligere i prosessen man kan fremskaffe pålitelige estimater jo større er nytteverdien (Gray, MacDonnell et al. 1999). Dilemmaet er at informasjonsgrunnlaget for å utføre estimeringen da blir spinklere og risikoen relatert til estimatene større.

Estimatene anvendes i forbindelse beslutningstagning, planlegging og oppfølging gjennom prosjektenes livssyklus. Forskningen på området har vist at Niels Bohrs utsagn om at "Det er vanskelig å prediktere, spesielt om fremtiden" i høyeste grad også gjelder for estimering av systemutviklingsprosjekter. Dette henger som Bohrs utsagn antyder, selvfølgelig sammen med at prediksjon i sin natur er vanskelig. Det er i tillegg en del karakteristika ved systemutvikling som er med på å gjøre utfordringene særlig store her. Fagfeltet er preget av en meget høy teknologisk utviklingstakt, en rask spredning til nye anvendelsesområder og høye krav til fleksibilitet og tempo i utviklingsprosjektene (Boehm and Sullivan 1999). Kombinasjonen av teknologiske, menneskelige og organisasjonsmessige aspekter knyttet til utviklingen av nye datasystemer kompliserer prosjektstyringen generelt og estimeringen spesielt. Dette gjør estimeringsarbeidet spesielt utfordrende, men samtidig også meget interessant.

På denne bakgrunn har estimering innen systemutvikling i løpet av de siste tretti år blitt viet betydelig forskningsfokus. Et karakteristisk trekk ved forskningen på området er at den har hatt et tungt innslag av empirisk tilnærming. Det er utarbeidet og videreutviklet en rekke metoder for estimering, og disse er i utstrakt grad blitt gjenstand for empiriske undersøkelser ofte i form av analyser mot ulike datasett av historisk gjennomførte prosjekter. Man har likevel ikke lyktes med å forske seg frem til en generell måte å estimere på som sikrer tilfredsstillende kvalitet på estimatene. Trass i lovende resultater ved anvendelse av spesifikke metoder i enkeltstudier, viser ulike empiriske undersøkelser som sammenligner treffsikkerheten på flere metoder sprikende og til dels motstridende resultater med hensyn til hvilke metode som gir de mest treffsikre estimatene (Jørgensen 2002).

Dette viser at de alternative eksisterende estimeringsmetodene har styrker og svakheter som slår forskjellig ut når metodene anvendes i ulike organisasjons- og prosjektkontekster. At metodene har sine styrker og svakheter er belyst i forskningen, og det er også argumentert for at disse forskjellene utgjør et potensiale man bør utnytte (Jørgensen and Sjøberg 2001). Lite forskning har derimot blitt gjort på forholdet mellom de enkelte metodene og den konteksten de anvendes i. Dette er et interessant felt som påvirker hele estimeringsprosessen.

## 1.2. Problemstilling

I dette hovedfagsarbeidet fokuseres det på en begrenset del av estimeringsprosessen, nemlig valg av estimeringsmetode i systemutviklingsprosjekter. Problemstillingen for arbeidet er å se på sammenhengen mellom den aktuelle prosjekt- og organisasjonskontekst og valget av estimeringsmetode.

Empiriske funn viser at det ikke eksisterer noen estimeringsmetode innen systemutvikling som er best uavhengig av kontekst. En gjennomgang av 15 ulike sammenligningsstudier viste at henholdsvis modellbaserte metoder og ekspertvurderinger var best i fem hver, mens det i de øvrige studiene ikke var noen klar tendens (Jørgensen 2002). Datasettene som anvendes i disse studiene er i utstrakt grad virksomhetsspesifikke eller bransjespesifikke samlinger av historisk gjennomførte prosjekter. Dette viser at valget av estimeringsmetode preges av kontekst.

Følgende sitat hentet fra Briand og Wiczorek sitt bidrag i boken Encyclopedia of software engineering viser det samme :

*“This overview article has shown that, although a number of alternatives exist to support the applications of software resource estimation, estimation methods all tend to have drawbacks and advantages. The appropriate selection of a method is therefore context dependent and this is why we have focused in this article on providing the readers with tools necessary to make an appropriate selection.” (Briand and Wiczorek 2002), side 58.*

Mitt arbeid har hatt som fokus å finne en tilnærming til valg av estimeringsmetode som best mulig sikrer at de kontekstuelle forhold får innflytelse på metodevalget. Selv om sitatet over antyder stor grad av sammenfall med Briand og Wiczoreks arbeid, er det etter min vurdering ikke tilfelle, da de nærmeste overser ekspertvurdering og bruker generelle evalueringskriterier. Min tilnærming skal bidra til mer bevissthet rundt valg blant alle tilgjengelig alternative estimeringsmetodene, og tilrettelegge for en mer prosjektilpasset estimeringsform.

Til mitt kjennskap er dette ikke omtalt som eget tema innen forskning på estimering, selv om ulike publikasjoner omtaler forhold som er med å påvirke metodevalget. Innen estimeringsforskningen har man i stor grad vært fokusert på å komme frem til en ”one-size fits all” estimeringsmetode med generell anvendelse. Det er brukt uforholdsmessig store forskningsressurser på å vise at enkelt metoder gir mer treffsikre estimater enn øvrige metoder, fremfor å se nærmere på potensialet i de alternative metodenes komplementære og supplerende egenskaper (Hughes 1996).

Resultatet er at man innen fagfeltet sitter med et stort antall alternative metoder, modeller og teknikker hvor man har empiri på at de gir god treffsikkerhet i estimatene i gitte sammenhenger. Det eksisterer derimot minimalt med hjelpemidler som estimeringsansvarlige kan anvende for utnytte denne akkumulerte kunnskapen når de skal bestemme seg for tilnærmingen til estimeringsarbeidet. For å oppnå en best mulig estimeringsprosess og et mest mulig treffsikkert estimat mener jeg det er viktig å velge en estimeringsmetode som passer i den gitte kontekst, og til dette trenger de ansvarlige støtte i form av retningslinjer eller rammeverk for å foreta et bevisst valg. Dette hovedfagsarbeidet ser nærmere på denne utfordrende problemstillingen.

### 1.3. Avgrensning

Arbeidet avgrenses i første rekke av at jeg ser på en snever del av den helhetlige estimeringsprosessen. Det er prosjektleder og eventuelt andre estimeringsansvarlige sitt valg av estimeringsmetode som diskuteres. Det er et mål å begrense diskusjonene i oppgaven til å ta opp forhold som angår selve metodevalget, og ikke komme inn på de mange øvrige utfordringene som finnes innen fagfeltet estimering. For å utforme et rammeverk for valg av metode trengs en innsikt i og forståelse for helheten og den samlede kompleksiteten innen fagfeltet. Gjennom arbeidet med oppgaven er mye ressurser brukt på å bygge opp en slik innsikt, og det ligger en fare i å miste fokus. Siden det er lite eksisterende forskning spesifikt på valg av metode, men en omfangsrik mengde av publikasjoner som inneholder elementer som jeg støtter meg på for å komme frem til et hjelpemiddel for metodevalg, blir utfordringen med å holde fokus på valgrelaterte temaer stor.

Rammeverket for å velge metode er ment å ha generell anvendelse innen systemutvikling. Det er skreddersydd for oppgaver som er relatert til utvikling av datasystemer, men det er ikke gjort noen avgrensning med hensyn til for eksempel hvilke type applikasjoner som utvikles. Derimot er det en avgrensning at jeg kun har tatt for meg oppgaver der prosjekt anvendes som arbeidsform. Det er ingen forutsetning at det er nyutvikling av applikasjoner, men eventuelle vedlikeholdsoppdrag som søker støtte i rammeverket bør være av en slik karakter og et slikt omfang at det gjennomføres i form av et prosjekt. Estimeringsforskningen har i en dominerende grad fokusert på nyutvikling fremfor også å omfavne vedlikeholdsoppdrag som utgjør svært omfangsrik del av all systemutvikling. Kun unntaksvis har vedlikeholdsoppgaver blitt viet oppmerksomhet fra forskere (Jørgensen 1995; Ferens 1999; Kitchenham, Pfleeger et al. 2002)

Relevansen av å benytte et slikt rammeverk avhenger av at det eksisterer en reell mulighet til å velge estimeringsmetode. Mitt inntrykk fra litteraturstudiene, egne erfaringer og uformelle henvendelser til profesjonelle systemutviklingsorganisasjoner, er at den enkelte prosjektleder i utstrakt grad selv kan velge tilnærming. Som følge av at valg av metode ikke har vært fremmet som eget tema finnes det få publikasjoner som har sett på dette forholdet. Heemstras kartlegging viser at det er spredning i hvilke metoder som velges, samtidig med at den synliggjør et lavt bevissthetsnivå rundt metodevalget (Heemstra 1992). En annen studie viser at ekspertvurderinger ble foretrukket som estimeringsmetoden i 104 av totalt 144 prosjekter i en organisasjon der man hadde åtte metoder å velge blant (Kitchenham, Pfleeger et al. 2002). Jeg har valgt å avgrense oppgaven til ikke å studere dette nærmere, men se på problemstillingen under forutsetning av at rammeverket er relevant i sammenhenger der valgmuligheten eksisterer.

### 1.4. Mål og målgruppe

Sluttmålet med det arbeid som her er igangsatt er etablering av et rammeverk som gir den nødvendige støtte ved valg av estimeringsmetode i systemutviklingsprosjekter. Hensikten er å tilby prosjektledere og estimeringsansvarlige et hjelpemiddel til anvendelse på det stadium i prosjektene der de skal bestemme seg for hvilke metode de skal støtte seg på under utarbeidelsen av estimatene.

På bakgrunn av at valg av estimeringsmetode har vært viet en så begrenset oppmerksomhet gjennom tretti års forskning innen estimering, er det et for ambisiøst mål for dette hovedfagsarbeidet å nå helt frem til sluttmålet. Arbeidet må betraktes som et det første viktige steg på veien der det er foreslått et rammeverk basert på en valgstrategi jeg oppfatter som meget lovende. Målet har vært å presentere en tilnærming som gir prosjektlederen den nødvendige oversikt over de alternative metodene og lansere et sett av vurderingskriterier som tilrettelegger for at den aktuelle organisasjonsmessige og prosjektmessige konteksten får prege metodevalget.

Målgruppen for dette utkastet til rammeverk er prosjektledere, estimeringsansvarlige og forskere med interesse innen dette området. Jeg har tilstrebet å komme frem til et utkast til rammeverk som er omfattende nok til at forskere kan gi konstruktive tilbakemeldinger. Samtidig bør prosjektledere med mer enn gjennomsnittlige interesse for utfordringene som ligger i estimeringsoppgaven, kunne gjennomgå rammeverket og anvende dette i konkrete prosjekter slik at de kan rapportere foreløpige erfaringer.

### 1.5. Metodevalg

En vitenskapelig arbeidsform fordrer at innsamling av data, behandling av informasjon og presentasjon av arbeidet gjøres på en systematisk måte (Jacobsen 2000). Det er i metodelitteratur innen naturvitenskapelig- og samfunnsforskning lansert et rikt utvalg av alternative metoder som skal støtte forskere i arbeidet med å tilstrebe dette.

Karakteristika ved problemstillingen som tas opp bør i stor grad være førende for metodevalget. Det er likevel ikke slik at det er riktige og gale metodevalg relatert til en problemstilling. Derimot vil ulike forskningsmetoder være mer og mindre hensiktsmessige. Det er viktig å redegjøre for metodevalget i et forskningsarbeid slik at andre forskere og interesserte kjenner tilnærmingen som ligger til grunn for arbeidet. Videre vil den som har gjennomført arbeidet også bli bedre i stand til å redegjøre for styrker og svakheter knyttet til resultatet. (Halvorsen 1993).

Dette hovedfagsarbeidet er det første større forskningsrelaterte arbeid jeg har gjennomført. Fremfor å gjøre en generell diskusjon om vitenskapelige metoder med tradisjonelle vurderinger rundt kvalitative kontra kvantitative metoder, beskrivelser av induktiv kontra deduktive tilnærminger og så videre, velger jeg her å beskrive og argumentere for den metodiske tilnærmingen jeg har hatt gjennom hele i mitt arbeid, altså også forut for at jeg definerte den endelige problemstillingen for arbeidet.

Min bakgrunn ved oppstart av hovedfaget var en rimelig bred teoretisk kunnskap om systemutvikling generelt. I tillegg hadde jeg praktisk erfaring rundt systemutvikling og prosjektledelse. Detalj kunnskapene innen estimering var begrenset til et overordnet inntrykk av viktigheten av å ha gode estimater, og kjennskap til COCOMO som estimeringsmetode. Dette er nok en bakgrunn som er karakteristisk for nettopp den målgruppen som trenger støtte innen området mitt arbeid belyser.

Problemstillingen for hovedfaget var ikke definert på forhånd. Følgende sitat hentet fra mitt essay levert i tilknytning til hovedfagsarbeidet i mars 2000 illustrerer innfallsvinkelen til det videre arbeid :

*”I mitt arbeid vil jeg studere enkelte eksisterende modeller mer i detalj for å danne meg et bedre bilde av styrker og svakheter ved modellene, og derigjennom kunne trekke slutninger om hvilke områder de ulike modellene bør benyttes innen. Videre vil det være sentralt å se på andre arbeider der man diskuterer kombinasjon av ulike modeller.”*

Denne innfallsvinkelen krevde et generelt innblikk innen estimering av systemutviklingsprosjekter, og de første månedene av hovedfagsarbeidet gikk derfor med til å sette seg inn i emnet gjennom litteraturstudier. Parallelt med dette ble en empirisk undersøkelse om valg og anvendelse estimeringsmetoder i et lite utvalg organisasjoner planlagt. Resultatet av disse innledende litteraturstudiene var at jeg fikk innblikk i et betydelig omfang av publisert forskning som i liten grad er systematisert slik at det kan anvendes av de som i praksis driver med estimering. Innen andre fagområder har artikler som (Mahmoud, DeRoeck et al. 1992) diskutert dette temaet, men innen systemutvikling fant jeg ingen tilsvarende publikasjoner. Parallelt med dette avdekket planleggingen av den empiriske undersøkelsen lite bevissthet rundt valg av estimeringsmetoder selv i organisasjoner som har sin virksomhet innen profesjonell systemutvikling. Dette ga meg inspirasjon til så vel den endelige problemstillingen som til valg av forskningsmetode i det videre arbeidet.

Temaet for det videre arbeidet ble avgrenset til å prøve å kombinere det behov praktikerne har for støtte i vurdering av estimeringsmetodene med den akkumulerte kunnskap som er representert i forskningen på estimering. Det store antall publikasjoner inneholder hver for seg empiri innen spesifikke områder, men satt inn i en sammenheng kunne dette bidra til utvidet kunnskap som kan være til nytte for så vel forskere som estimeringsansvarlige i prosjekter.

Problemstillingen ble fokusert på å gjøre kontekstavhengige valg av estimeringsmetode. Planen om å gjennomføre en empirisk undersøkelse i en gitt kontekst ble forlatt. Fremfor å tilføye enda et empirisk funn til mengden av det som allerede er publisert, valgte jeg å fordype meg i litteraturstudier som forskningsmetode. Til grunn for mitt arbeid ligger en induktiv tilnærming, der jeg på bakgrunn av tidligere publisert forskning har jobbet med å utdype kunnskapen om valg av estimeringsmetode. Jeg har hatt en kvalitativ metodebruk der tolkning av sammenhenger som påvirker valgssituasjonen har vært sentralt. Jobben har gått ut på å identifisere og systematisere relevante funn i publisert forskning, og på grunnlag av dette formulere et rammeverk for valg av estimeringsmetoder.

Litteraturstudiene innen estimerings av systemutvikling ga liten støtte i forhold til strategisk tilnærming til valg av metode. Her måtte jeg derfor å hente inspirasjon fra publikasjoner som tar for seg prediksjoner innen andre fagfelt. Et annet område jeg måtte foreta søk i litteratur utenfor fagfeltet var innen ekspertvurderinger, da forskningen ikke står i forhold til den store utbredelsen av denne estimeringstilnærmingen har innen systemutvikling.

Innen rammene av et hovedfagsarbeid har arbeidet med å foreta en litteraturstudie gitt god totalforståelse for feltet, og muliggjort lansering av et forslag til rammeverk for valg av metode. Det har vært riktig å prioritere å komme lengst mulig i systematiseringen innen dette feltet, fremfor å supplere litteraturstudiene med øvrige forskningsmetoder. Det er her viktig å ta med i betraktningen den forskningstilnærming



som er representert i de publikasjoner som er gjennomgått. De mange publikasjonene har en sterk empiriske fokus i form av post-mortem analyser, survey's, case-studier og også enkelte eksempler på eksperimentell tilnærming, og dette fører til at mine litteraturstudier på denne publisert forskning får et empirisk fundament.

I tillegg til å støtte meg på så vel spesifikke funn og generelle inntrykk fra litteraturstudiene, har det vært nødvendig å bruke egne erfaringer og sunn fornuft i arbeidet. Satt inn i en sammenheng der dette er med å supplere og utfylle en systematisk vitenskapelig tilnærming, er en slik tilnærming forsvarlig. Det er likevel med på å understreke at rammeverket som er lansert ikke representerer et ferdigstilt og empirisk bevist hjelpemiddel, men heller et lovende utgangspunkt for ytterligere studier.

## 1.6. Oppgavens struktur

Oppgavens struktur er bygd opp rundt det rammeverk som jeg gjennom funn i mine litteraturstudier har lansert som et lovende forslag til løsning på den definerte problemstilling. Etter først å plassere problemstilling jeg tar opp innen fagfeltet, går jeg gjennom de tre sentrale elementene i mitt arbeid, nemlig valgstrategi, metodealternativene og de kontekst-fokuserte valgkriteriene. I gjennomgangen av disse vises det til relevante funn gjort i litteraturstudiet som er med på å underbygge tilnærmingen. Avslutningsvis vurderes helheten i rammeverket før hovedfagsarbeidet oppsummeres. Denne strukturen har gitt følgende kapittelinndeling :

### **Kapittel 2 : Sentrale trekk fra litteraturstudien**

Her presenteres enkelte grunnleggende trekk ved estimeringsaktiviteten i systemutviklingsprosjekter som i særlig grad er med på å sette rammene for det senere valget av metode. For å gi et bakteppe for den påfølgende presentasjon av rammeverket for metodevalg, er det gitt et kort historisk tilbakeblikk på hvordan man har betraktet den helhetlige estimeringsprosessen innen forskningsfeltet. Problemstillingen jeg tar opp blir plassert inn i den totale estimeringsprosessen. Metodenes ulike måter å kombinere analytisk og intuitiv tilnærming på diskuteres, og aktualiteten av å kombinere ulike estimeringsmetoder innen et prosjekt belyses.

### **Kapittel 3 : Strategi for valg av estimeringsmetode**

Her beskrives og drøftes ulike strategier for å valg estimeringsmetode. I mangel på forskning på dette feltet innen systemutvikling, diskuteres relevansen av de alternative strategier for valg av prediksjonsmetode som er presentert av Armstrong (Armstrong 2001) for systemutviklingsprosjekter. Det konkluderes med at *Strukturert Vurdering* er en lovende strategi, og denne velges som basis for det videre arbeid med å utarbeide rammeverket.

### **Kapittel 4. Kartlegging av alternative estimeringsmetoder**

I en *Strukturert Vurdering* er det en forutsetning at man kjenner de aktuelle alternative estimeringsmetodene man har å velge mellom. Her kategoriseres estimeringsmetodene innen fagfeltet. Inndeling i hovedkategoriene Modellbasert estimering, Analogibasert estimering og Ekspertvurderinger begrunnes, og det gis en mer detaljert sub-kategorisering av disse. Forskningen innen alternative tilnærminger til ekspertvurdering er svært mangelfull innen vårt fagfelt og vies derfor spesiell fokus.

**Kapittel 5. Kriterier for å foreta et kontekstavhengig metodevalg**

Med basis i litteraturstudiene lanserer jeg et sett av kriterier som skal sikre at valget av estimeringsmetode ivaretar den aktuelle kontekst. Hvert av kriteriene beskrives og relateres til problemstillingen. Deretter vurderes de konsekvenser dette kriteriet har for metodevalget, før gjennomgangen avsluttes med å referere til relevante publikasjoner der dette er funnet.

**Kapittel 6. Anvendelse av rammeverket**

Her antyder jeg hvordan jeg ser for meg at rammeverket kan bli anvendt i praksis. Jeg har forsøk å se det fra de fremtidige brukernes perspektiv.

**Kapittel 7. Konklusjon og forslag til videreføring**

Det gis en redegjørelse for de viktigste trekkene i det mitt arbeid har resultert i og et forslag til videreføring av arbeidet.

## 2. SENTRALE TREKK FRA LITTERATURSTUDIEN

Jeg har jobbet med å skaffe meg innblikk i forhold som berører problemstillingen om valg av estimeringsmetode fra et bredt spekter av litterære kilder. Denne studien danner basisen for presentasjonen av en tilnærming til å løse problemstillingen i oppgaven. Mitt perspektiv preges av at jeg mener vi bør mye bli bedre til å anvende akkumulert kunnskap fra forskningen også i forbindelse med metodevalg innen estimering.

Som en innledning til lanseringen av rammeverket gir dette kapitlet en presentasjon av noen utvalgte grunnleggende trekk jeg fant i litteraturstudiene med stor innvirkning på denne tematikken. Det gis et innblikk i forskernes betraktninger rundt sentrale elementer i estimeringsprosessen gjennom de siste tretti år. Valget av estimeringsmetode er en liten, men viktig del av denne prosessen, og en forståelse av utviklingen på feltet vil danne et gunstig bakteppe for den videre diskusjon. Hvilke nivå i organisasjoner valgene gjøres på blir kort kommentert. Kapitlet avsluttes med å se på hvordan metodene kombinerer analyse og intuisjon, før potensialet som ligger i å kombinere flere metoder innen et prosjekt diskuteres.

### 2.1. Estimering – en støtteaktivitet ved utvikling av datasystemer

Et estimat er i denne sammenheng en prediksjon på hvor store ressurser man mener er nødvendig å bruke for å gjennomføre et forestående systemutviklingsprosjekt. DeMarco definerte et estimat til å være en prediksjon som like sannsynlig er over som under det endelige resultat (DeMarco 1982). Målet med estimering er å gi en prediksjon som er god nok til å fatte de rette beslutninger, ikke at man skal treffe nøyaktig. Metodenes treffsikkerhet har som jeg vil komme tilbake til, tradisjonelt vært det viktigste kriteriet for valg av metode. Et interessant trekk ved estimering, er en generell tendens mot å lage for snevre intervaller på estimatene.

*”Estimates showed a varying degree of overoptimism and overpessimism but all were much too tight, with almost half of actual outcomes falling in the 1 % tails of estimated distribution”. (Connolly and Dean 1997) side 1029.*

Temaet belyses nærmere i (Jørgensen and Moløkken 2002) som ser på ulike strategier for å øke treffsikkerheten på disse intervallene. Denne svakheten ved metodene er med på å underbygge at det ikke er gunstig å basere et valg av estimeringsmetoder kun på graden av treffsikkerhet. Det er for stor usikkerhet forbundet med dette kriteriet innen fagfeltet..

Etter min oppfatning er en annen generell svakhet ved metodene innen fagfeltet er at de er altfor lite eksplisitte i å uttrykke hvilke utviklingsressurser som inngår i estimatene. En del generelle utviklingsaktiviteter vil alltid inngå, men for den som skal bruke estimatene er det viktig å vite om for eksempel administrative oppgavene eller brukernes medvirkning også inngår i estimatene. Avklaringer av hva som inngår burde vært mer fokusert i publikasjoner som tar for seg spesifikke metoder. For de som skal velge metode stilles det derfor krav til at de selv gjør seg opp en klar oppfatning av hvilke

ressurser de ønsker et estimat på, og lar dette prege så vel metodevalg som metodeanvendelse.

Alle som arbeider innen fagfeltet estimering enten det er innen forskning eller i praktisk anvendelse, må forholde seg til det faktum at estimering ikke er av kjerneaktivitetene ved systemutvikling. Med kjerneaktiviteter mener jeg her de aktiviteter som må gjennomføres for få være i stand til å ferdigstille en programvare. Generiske oppgaver som kravspesifisering, ulike former for design, koding og testing vil alltid inngå i et helhetlig systemutviklingsprosjekt. Øvrige aktiviteter er støtteaktiviteter man kan velge å prioritere i større eller mindre grad, og de vil i mange sammenhenger ikke omfattes av samme faglige prestisje som de rent utviklingsorienterte aktivitetene. Prioriteringen av hvilke støtteaktiviteter man vil fokusere og omfanget av ressurser man er villig til å bruke på den enkelte, vil variere i ulike organisasjoner og mellom de enkelte prosjektene. Dette vil innvirke på metodevalget da ulike metoder har ulike krav til omfang på estimeringsressurser. Kvalitetssikring, prosessforbedring, risikovurdering, planlegging, ledelse, estimering, oppfølging og konfigurasjonsstyring er alle eksempler på støtteaktiviteter som bidrar til å forbedre en utviklingsprosess.

Selv om man historisk sett har dårlig resultater å vise til når det gjelder treffsikkerhet på estimater innen systemutvikling, medvirker krav fra omgivelsene og estimatenes store potensialet i form av kritisk styringsinformasjon til at aktiviteten normalt gjennomføres i systemutviklingsprosjekter. Ressursene man er villige til å sette på oppgaven er derimot ofte svært begrensede. Tilnærminger som med begrenset ressursbruk og lave krav til forkunnskaper gir treffsikre estimater har derfor vært forskernes siktemål. Det generelle inntrykket jeg sitter med etter å ha satt meg inn i ulike estimeringsmetoder er at de har et realistisk utgangspunkt med hensyn til ressursbruk. Måten de tilstreber en lav ressursbruk på er derimot høyst forskjellige. Dette angår metodevalget og belyses nærmere senere i oppgaven.

For å være et redskap systemutviklerne kan anvende i praksis, må estimeringsmetodene og estimatene ha en del kjennetegn som gjør dem attraktive å anvende. Etter å ha gjennomgått et stort antall publikasjoner om temaet, har jeg kommet frem til at essensen i kravene som stilles til estimeringsmetodene etter min vurdering kan sammenfattes i følgende punkter :

Metodene bør :

- raskt og med en begrenset ressursbruk gi estimater med en treffsikkerhet som tilfredsstiller behovet
- kunne anvendes på det tidspunkt i utviklingsprosessen man trenger estimatene
- kunne nyttiggjøre seg den informasjon som er tilgjengelig på estimeringstidspunktet
- være lettforståelig og egnet til å skape eierskap til estimatene
- kunne tilpasses lokale forhold

Når jeg senere presenterer et rammeverk for å velge estimeringsmetode, bør graden av oppfyllelse av overstående punkter klart influere på metodevalget. Å tilfredsstille alle kjennetegnene har gjennom tretti års forskning vist seg urealistisk. Et hovedpoeng i min tilnærming er at metodene i ulik grad tilfredsstiller de enkelte kjennetegnene, og at nettopp det forholdet må være med å avgjøre hvilken metode man velger.

### 2.2. Begrepsbruk

#### 2.2.1. Estimeringsmetode

Før jeg går videre med å diskutere valg av estimeringsmetode, er det viktig å presisere at det ikke eksisterer noen generell og allment akseptert felles oppfatning av hva som ligger i begrepet ”estimeringsmetode”. I mitt arbeid har jeg funnet det hensiktsmessig å benytte den definisjonen av begrepet som Briand og Wieczorek gir i sitt rammeverk:

*”It consists of either (1) one or several models, possibly a modelling method and a model together with a method to apply the model(s), or (2) one or more estimation techniques. This distinction is made because some estimation methods do not include any explicit resource modelling. The estimates are solely based on expert judgement.” (Briand and Wieczorek 2002) side 25.*

Av denne fremgår det at modellene og teknikkene er elementer som inngår i en metode. Over tid har disse begrepene blitt benyttet om hverandre, noe man vil se der jeg gjengir konkrete eksempler fra publikasjoner.

#### 2.2.2. Estimattyper

Ved siden av estimater for ressursbehov i form av utviklingsinnsats og utviklingskostnad, estimerer man også varigheten på utviklingsprosjekter. Innen det temaet dette hovedfagsarbeidet tar opp er alle disse estimattypene relevante, men estimering av utviklingsinnsatsen er den type estimat jeg har fokusert på. Dette speiler den oppmerksomhet de enkelte estimatene er gitt innen forskningen, og man mener at så vel det meste av kostnadene og som varigheten på et prosjekt kan utledes fra et estimat på utviklingsinnsatsen. I resten av oppgaven er det estimering av innsatsbehov, i engelskspråklige publikasjoner kalt ”effort estimate”, som ligger i begrepet estimat der ikke annet er spesifisert.

### 2.3. Valg av estimeringsmetode – et lite omtalt tema i litteraturen

En gjennomgang av sentrale lærebøker innen systemutvikling og de publiserte artiklene som er samlet i BEST-biblioteket <sup>1</sup> viser at ”valg av estimeringsmetode” så langt jeg har kunnet finne ikke har vært gjenstand for nærmere diskusjon og forskning. Informasjonssøk gjort i ulike forskningsdatabaser som INSPEC og Science Direct samt utstrakte uformelle søk i tilgjengelig materiale på internett forsterker inntrykket av at dette er et lite berørt tema.

Litteraturstudiet har vist at man innen forskningsfeltet har vært fokusert på å utvikle og forbedre algoritmiske modeller for estimering av kostnad og innsats. Som et alternativ har mindre forskningsmiljøer jobbet med en analogibasert tilnærming, mens den i praksis svært utbredte ekspertestimeringen inntil nylig har vært viet minimal

---

<sup>1</sup> En samling publiserte artikler innen estimering av systemutviklingsprosjekter initiert av BEST-prosjektet, ajourføres nå i form av en ENDNOTE-base av forskningsorganisasjonen simula research laboratory. Denne ENDNOTE-basen har vært bruk som utgangspunkt for referansehåndtering i dette hovedfagsarbeidet.

oppmerksomhet fra akademiske miljøer. Det er overraskende at valg mellom alternative estimeringsmetoder i liten grad har vært gjenstand for forskning. Enda mer oppsiktsvekkende er det at muligheten for faktisk å velge estimeringsmetode i liten grad trekkes inn som et interessant perspektiv i publikasjoner som dekker estimeringsprosessen som helhet. Under oppsummerer jeg noen sentrale trekk jeg har funnet rundt temaet innen ulike kategorier av publikasjoner jeg har studert.

### **2.3.1. Lærebøker innen Software Engineering**

Publiserte forskningsartikler har utgjort den vesentlige informasjonskilden i hovedfagsarbeidet, men for å få et innblikk i den introduksjon de generelle lærebøkene innen systemutvikling gir om temaet estimering, er også ti utvalgte bøker studert. For mange som utfører estimering i praksis består deres teoretiske bakgrunn innen emnet av at de har lest om temaet i denne type lærebok, noe som er med på å gjøre grunnlaget for et bevisst og reflektert valg spinkelt.

Lærebøkene innen systemutvikling har bred fokus og tabellen under viser at innslaget av estimering er relativt beskjedent. Bøkene legger i sin dekning av estimering generelt vekt på å beskrive viktighet av og kompleksiteten i å estimere systemutviklingsoppdrag. Aktiviteten plasseres i den delen av boken som tar for seg prosjektstyring av systemutviklingsprosjekter, og estimatenes rolle i forbindelse med beslutningstaking og ressursplanlegging belyses på en rimelig balansert måte.

Et generelt trekk ved bøkene er at de etter denne innledende diskusjonen presenterer en enkelt metode for hvordan man kan gjennomføre estimeringen. I det utvalg av bøker som her ble gjennomgått, var ulike versjoner av COCOMO den klart mest omtalte metoden. Til tross for at et stort forskningsmaterieell viser at denne metoden på ingen måte er blant de metodene som kan vise til best treffsikkerhet (Briand and Wiczorek 2002), har den en stor utbredelse i lærebøkene, og blir således den enkeltmetoden som flest utviklere har blitt introdusert for.

## 2. SENTRALE TREKK FRA LITTERATURSTUDIEN

Tittel	Forfatter, år	Dekning av estimering	Omtalte tilnærminger	Omtalte enkeltmetoder
Software Engineering	R.Pressman, 1997	30 av 885 sider	Dekomponering,	COCOMO (FP og LOC)
Software Engineering	S.P.Pfleeger, 2001	26 av 659 sider	Ekspert vurdering, Algoritmiske, Nevrale nett	COCOMO
Systemutvikling	E.Andersen, 1994	0 av 436 sider		
Systemtutvikling - fra kjernen og ut, fra skallet og inn	G.Skagestein 2002	16 av 477 sider	Eksplisitt Analogi, Fremskrivning, Modellkalkyler, Ekspertvurdering	FP
Software Engineering	I. Sommerville	24 av 693 sider	Algoritmiske modeller,Analogi, Ekspertvurdering,	COCOMO
System Analysis and Design	A.Dennis m.fl, 2003	20 av 535 sider	Timeboxing	FP, COCOMO
Fundamentals of Software Engineering	C. Ghezzi m.fl., 2003	17 av 604 sider	WBS	FP, LOC, COCOMO
The Engineering og Software	D.Hamlet m.fl, 2001	2 av 494 sider		COCOMO
Software Engineering	D. Bell	4 av 470 sider		COCOMO
Project-based Software Engineering	E.Stiller, C.LeBlanc 2002	4 av 382 sider		FP

*Tabell 2.1. Dekning av estimering i lærebøker om "Software Engineering"*

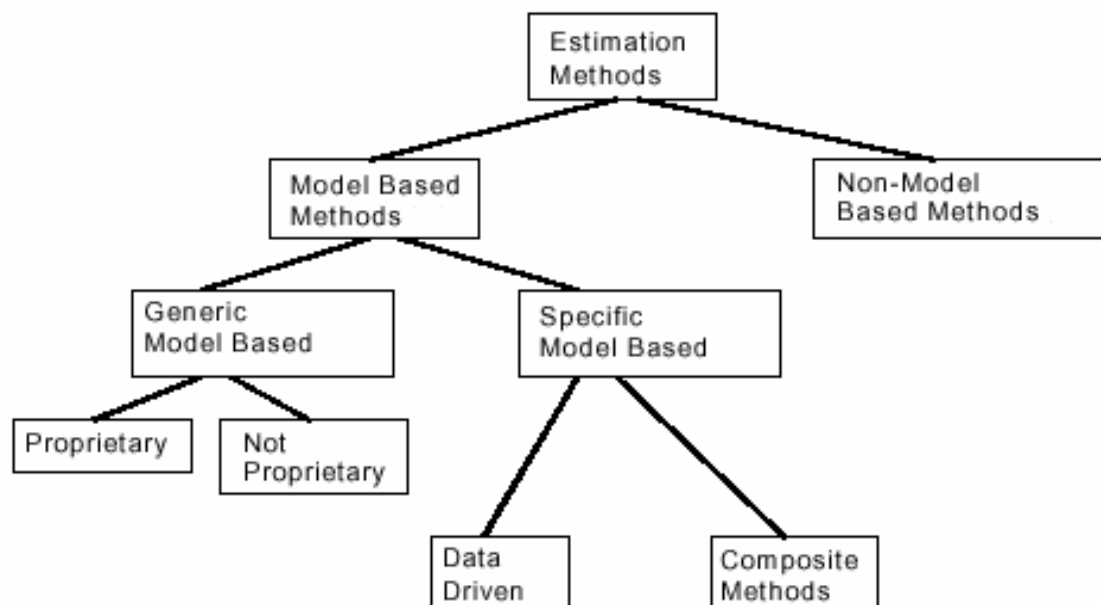
Gjennomgangen viste at det relative styrkeforholdet mellom beskrivelsene av ulike estimeringsmetoder tenderer mot en kraftig overrepresentativitet av modellbaserte metoder. Dette viser samme skjevhet som i den publiserte forskningen. Dette er uheldig i forhold til at et metodevalg bør preges av at man har en god oversikt over de ulike metoder og deres styrker og svakheter. Flere av bøkene nevner overhodet ikke ekspertvurderinger, og ingen av bøkene gir noen dypere beskrivelse av hva som ligger i denne tilnærmingen.

Utover de generelle lærebøkene finnes det også bøker som tar for seg temaet kostnads- og innsatsestimering av systemutviklingsprosjekter spesielt. Enkelte bøker går her svært grundig inn på å presentere enkelte modellbaserte metoder. Boehm gjennomgår i sine bøker (Boehm 1984; Boehm, Abts et al. 2000) COCOMO metoden, mens Londeix (Londeix 1987) fokuserer mest på Putnam-baserte estimeringsmetoder, men omtaler også COCOMO. På den andre siden finner man bøker som ser det hele fra et rent praktisk perspektiv og på et overordnet nivå diskuterer estimatenes rolle i planleggingen (Coombs 2003). Jeg har ikke vært i stand til å finne bøker som på en grundig, objektiv og balansert måte diskuterer de ulike kategoriene metoder som finnes, og som ser nærmere på valget mellom dem. Jeg har ikke funnet noen bøker som går spesifikt på bruk av ekspertvurderinger innen fagfeltet.

### 2.3.2. Publikasjoner på valg av estimeringsmetode

Jeg har funnet to publikasjoner innen fagfeltet systemutvikling der temaet valg av estimeringsmetode behandles i en viss utstrekning. Den første er (Briand and Wieczorek 2002). De gir i hovedsak en beskrivelse av et antall modellbaserte estimeringsmetoder for så å gjøre en systematisk sammenligning av disse ut fra et evalueringsrammeverk. Sammenligningen gir et inntrykk av metodenes styrker og svakheter, men fremstillingen blir meget skjev ettersom ekspert vurderinger gis en svært overfladisk behandling.

De gir følgende klassifisering av eksisterende estimeringsmetoder. Denne kommer jeg tilbake til i kapitel 4.



Figur 2.1. Classification of resource estimation methods (Briand and Wieczorek 2002), side 5.

De ser avslutningsvis i sitt arbeid på organisasjonsspesifikke og tekniske forhold som de argumenterer for at bør påvirke valg av estimeringsmetode. Deres grunnholdning er likevel at valget av estimeringsmetode bør være empirisk begrunnet.

*"Researchers need to consistently compare estimation methods across many data sets in order to be able to define guidelines about what kind of models to use for a given environment." (Briand and Wieczorek 2002) side 56.*

Dette skiller seg fra min tilnærming. Jeg mener at man heller bør benytte akkumulerte funn fra forskningen som bakgrunnsmateriale. Utøverne bør selv gjøre en subjektiv faglige vurdering og forstå valget av metode i det enkelte prosjekt. Forskningen bør støtte opp med et hjelpemiddel som muliggjør dette ved hjelp av liten tidsbruk.



Også (Walkerden and Jeffery 1997) skisserer et rammeverk for valg av estimeringsmetode, men de har bevisst valgt å kutte Ekspertvurderinger helt ut. Ekspertvurderinger omtales her som en metode som vanskelig kan karakteriseres, og de mener videre at siden den alltid velges hvis eksperter er tilgjengelig så er det liten hensikt å omtale den i et rammeverk for valg av metode. Deres rammeverk omtaler fire ulike prediksjonsmetoder : Empiriske, Analogiske, Teoretiske og Heuristiske.

De lanserer et deskriptivt rammeverk, der målinger står svært sentralt. Inspirert av GQM (Goal Quality Metrics) og QIP (Quality Improvement Process) fokuserer de på målinger og forbedringer, fremfor valgsituasjonens iboende kompleksitet. Etter min opfatning virker de mer fokusert på å tolke og analysere eksisterende forskning på fagfeltet, enn på å hjelpe utøveren som står i valgsituasjonen. Dette blir helt forskjellig fra min innfallsvinkel, der fokus har vært å bistå utøveren ved å se nærmere på hva forskningen kan bidra med innen veiledning ved valg av estimeringsmetode.

### 2.3.3. Forskning som sammenligner metoder

Det finnes et betydelig antall kilder der man gjør sammenligninger av treffsikkerheten til de ulike estimeringsmetodene. Her rangeres modellene i forhold til hverandre, men det gis sjelden noen analyse av i hvilke kontekster de ulike metodene har sine styrker og svakheter. Disse sammenligningene gir ofte ulike og til dels motsigende konklusjoner, noe som jeg mener forklares med at datasettene man bruker kan representere ulike kontekster. Som tidligere kommentert gir (Jørgensen 2002) av oppsummering på 15 ulike sammenligningsstudier.

Lederer og Prasad mener at det er en mangel på objektive studier som sammenligner metoder. De viser til at mange av studiene er gjort av forskere som selv har lansert en av metodene, og at det da er fare for at disse blir selvoppfyllende og tenderer mot å støtte den metode som forskeren har fremmet. En annen stor svakhet er at studiene i mange sammenhenger blir urealistiske, da de gjør sine sammenligninger på grunnlag av allerede avsluttede prosjekter, og at metodenes styrkeforhold kan bli fordreid av denne urealistiske settingen (Lederer and Prasad 1992).

Utvalget av metoder man sammenligner sin egen metode kan vanskelig bli objektiv når forskerne som selv velger ut hvilke metoder man trekker inn. Når man ser publikasjoner som sammenligner sin metode mot ukalibrerte modellbaserte metoder til tross for at det er bred enighet om at disse modellene må kalibreres før de kan gi tilfredsstillende resultater, styrkes synet på at man ikke alltid skal ha tiltro til at sammenligningene er objektive.

Mange forskere bruker de svake resultatene som mange av disse sammenligningsstudiene viser, som argumentasjon for stadig å forske på utvikling av nye metoder. Følgende sitat fra eksemplifiserer dette :

*“It is stressed that there are a need for a new model because the existing models give inaccurate, inconsistent and unreliable predictions. Additionally they are based on either inappropriate variables or variables that can not be measured at the beginning of the development process. Finally, existing models do not support the planning of individual stages of the development process, but only try to make predictions about the project*

*development process as a whole.” (Chatzoglou and Macaulay 1996) side 173*

I mitt arbeid tar jeg et helt annet perspektiv, og ser heller på hvordan vi kan bli bedre til å benytte de allerede lanserte metodene innen fagfeltet. Det er tross alt gjort omfattende forskning som viser at metodene fungerer innen gitte rammer. Det gjelder bare å finne ut av hvilken av metodene som er mest egnet for bruk i den kontekst man står overfor i det aktuelle prosjektet.

### **2.3.4. Forskning på enkeltmetoder**

I publikasjoner som tar for seg enkeltmetoder brukes mye ressurser på å forklare den nye eller tilpasningene av den gamle estimeringsmetoden. Videre er en utbredt forskningstilnærming at man foretar statistiske analyser for å styrke at den aktuelle metoden gjennom å vise at den gir god treffsikkerhet i estimatene relativt til andre metoder. Det ser ut til at de fleste forskere som lanserer nye/forbedrede modeller mener disse er generelt anvendelige, da det er sjelden at man spesifiserer hvilke kontekst teknikken passer i. I de siste årene har riktignok kommet en del estimeringsmetoder som spesifikt tar for seg estimering av WEB-applikasjoner (Reifer 2000; Ruhe, Jeffery et al. 2003).

Ofte er empirien i disse arbeidene basert på anvendelse av metoden mot et spesifikt datasett. Estimeringen gjøres i de fleste tilfeller etter at prosjektene er gjennomført, og viser ofte relativt små, men statistisk holdbare forskjeller i treffsikkerheten til de ulike estimeringsmetodene. Ofte trekkes COCOMO inn som en metode man sjekker den nye metoden opp mot. Det kan virke som et etablert kriterium for å få aksept for en nylansert metode er blitt at den i alle fall må gi bedre treffsikkerhet enn COCOMO. Det kan stilles spørsmål ved en slik tilnærming, da de som utøver estimering i svært liten grad anvender COCOMO. Det ville etter min mening vært naturlig å sammenligne nylanserte metoder med ekspertvurderinger, da denne har svært stor utbredelse i praksis, og på mange måter er den metoden forskerne har som målsetting å utkonkurrere.

### **2.4. Metodevalgets plassering i estimeringsprosessen**

Estimering av innsatsbehovet i et systemutviklingsprosjekt er en oppgave som følger et utviklingsprosjekt gjennom dets livsløp, der det er en gjensidig påvirkning mellom estimeringsarbeidet og det utviklingsarbeidet som foregår (Abdel-Hamid and Madnik 1983). En god estimeringsprosess er preget av intensiv aktivitet tidlig i prosjektet, men også en stor grad av ajourhold og oppfølging utover i forløpet. Valget av estimeringsmetode bør etter min mening utgjøre en liten, men sentral del av estimeringsprosessen. I forhold til det videre arbeidet er viktig å plassere valget av estimeringsmetode som en aktivitet tidlig i estimeringsprosessen.

Som et bakteppe for den videre diskusjon velger jeg å trekke frem noen eksempler fra litteraturen på hvordan estimeringsprosessen er blitt beskrevet. Dette er viktig for å vise at valg av metode i liten grad er blitt betraktet som noe sentralt tema historisk sett.

### 2.4.1. Estimeringsprosesser beskrevet på 80-tallet

I den ofte refererte boka "Software Engineering Economics" lanserte Barry Boehm en estimeringsprosess bestående av følgende syv steg (Boehm 1984):

- |  |
|--|
| <ol style="list-style-type: none"><li>1. Establish objectives</li><li>2. Plan for required data and resources</li><li>3. Pin down software requirements</li><li>4. Work out as much detail as feasible</li><li>5. Use several independent techniques and sources</li><li>6. Compare and iterate estimates</li><li>7. Follow up</li></ol> |
|--|

*Tabell 2.2. Boehms klassifisering*

Ser vi denne i sammenheng med Boehm og Wolvertons artikkel (Boehm and Wolverton 1980) som belyser kravene man bør stille til en god estimeringsmodell, var denne prosessbeskrivelsen på mange måter forut for sin tid. Den preges av et bredt perspektiv på hva som bør inngå i en helhetlig estimeringsprosess. Bevissthet i hva man ville med estimatene skulle legges til grunn for hvordan man i neste omgang la opp hele estimeringsarbeidet. Man så også allerede her fordelene med at alternative metoder kan komplementere hverandre, og lanserte i det femte steget fordelene med å kombinere ulike tilnærminger.

Problemet med disse arbeidene var nok i stor grad mangelen på tilgjengelig metoder både med hensyn til bredde i alternativene og med hensyn til kvalitet på de etablerte metodene. Selv om prosessrammeverket virket fornuftig, ble det vanskelig å etterleve i praksis. Selv om dette gjennom historien stadig har vært referert som arketyper på en god estimeringsprosess av andre forskere, har evnen til å etterleve prosessen uteblitt, og dermed også utbredelsen av en slik tilnærming til estimering.

Et paradoks i forbindelse med denne estimeringsprosessen er Boehm's lansering av COCOMO som en estimeringsmodell i samme bok. Etter først å ha beskrevet de mangesidede og komplekse problemstillingene innen estimering, ble en algoritmisk estimeringsmodell med sterk forankring av en tidlig prediksjon på størrelsen av programvaren lasert av Boehm som en løsning på hvordan estimere. Slik jeg tolker senere forskning på området, har nettopp COCOMO modellen og den tilnærmingen denne representerer vært en inspirasjon for forskning innen modell-basert estimering. Dette virker til å ha fjernet fokus fra de mer helhetlige og komplekse problemstillingene. Detaljerte forskning rundt å finne generelle modeller for estimering ble den toneangivende trenden innen fagfeltet. Lansering av ulike algoritmiske modeller med en tilhørende bevisførsel på modellenes treffsikkerhet basert på historiske datasett kom til å prege publikasjonene.

Et annet eksempel på modell over den totale estimeringsprosessen fra denne tiden er DeMarco's single factor cost modell (DeMarco 1982).

- |   |
|---|
| <ol style="list-style-type: none"><li>1. Select the cost components to be modelled</li><li>2. Select the measures from which to predict the component and total cost</li><li>3. Develop the related single-factor cost models, based on existing data</li><li>4. Estimate cost and standard error of the estimate for each component model</li><li>5. Adjust the estimates based on differences between the current and past projects</li><li>6. Calculate the total cost from the component costs</li><li>7. Calculate a total estimating error</li><li>8. Use a time sensitive cost model to constrain the estimate of total project effort</li></ol> |
|---|

*Tabell 2.3. De Marco's Single Factor Cost Model*

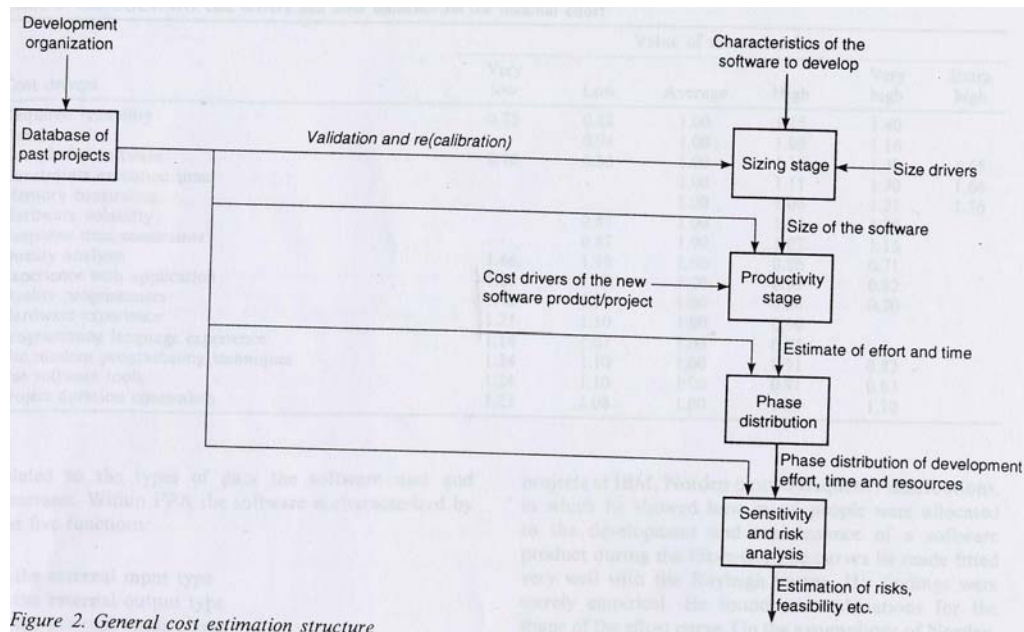
Et interessant trekk ved denne prosessen er at tilpasning, kalibrering eller utvikling av selve estimeringsmodellen inngår som en sentral del av estimeringen. Tilnærmingen til estimering er gitt av den modellen man anvender, men allerede her så man at det var et behov for å tilpasse modellene som ble brukt til den lokale kontekst man hadde i det enkelte utviklingsprosjekt.

### **2.4.2. Estimeringsprosesser beskrevet på 1990-tallet**

Estimeringsprosessene var på denne tiden sterkt to-fase delt, der man i første fase fokuserte på å estimere størrelsen på programvaren som skulle utvikles. Deretter jobbet man med å estimere produktiviteten til utviklingsgruppen, slik at man til slutt kunne gi estimerer for innsatsbehovet i antall månedsverk og varigheten på prosjektet i kalendertid. Jeg tolker denne utviklingen som et svar på den sterke posisjonen modellbasert estimering hadde innen forskningen. Nye tilnærminger til å prediktere størrelse og produktivitet avstedkom et utvidet spekter av estimeringsmetoder, samtidig som man videreutviklet og forbedret de eksisterende.

Under følger to eksempler fra forskningspublikasjoner som synliggjør fokuseringen. Det første eksemplet er Heemstra's "General cost estimation structure". Denne prosessbeskrivelsen viser godt de karakteristiske trekk for hvordan forskningen betraktet estimeringsprosessen i denne perioden.

## 2. SENTRALE TREKK FRA LITTERATURSTUDIEN



Figur 2.2. Modell av Heemstra (Heemstra 1992).

Første fase går ut på å gi en prediksjon på størrelsen på programvaren. Denne ble typisk angitt i enten antall kodelinjer (LOC) eller i funksjonspoenger (FP). Vi ser at karakteristika for det aktuelle prosjektet ble brukt som informasjonskilde, men det var ingen forutgående prosess som gikk på å vurdere ulike metoder. Det som inngår er en validering av estimeringsmodeller som ikke tidligere er benyttet i organisasjonen samt en kalibrering / rekalkibrering av modellen ut fra en lokal database med gjennomførte prosjekter for at den skal bli best mulig tilpasset lokale forhold.

Det andre eksemplet er Software Engineering Process Office (SEPO) 1999 versjon 2.2.

1. Estimat Size
2. Estimate Cost and Effort
3. Estimate Schedule
4. Estimate Critical Computer Resources
5. Assess Risks
6. Inspect / Approve
7. Track and Report Estimates
8. Measure and Improve Process

Tabell 2.4. SEPO prosessen (versjon 2.2. hentet fra <http://sepo.spawar.navy.mil/sepo>)

Begge eksemplene viser den sterke fokus på en innledende estimering av størrelse med en påfølgende estimering av innsats og varighet. SEPO's prosess viser en positiv utvikling i retning av at man etter hvert setter mer fokus på å følge opp estimeringen utover i prosjektet. Ajourføring av estimatene utover i prosjektet og fokus på å måle hvor god man har vært til å estimere viser at man her begynner å trekke inn læringsaspektet i estimeringsprosessen.

### 2.4.3. Nyere estimeringsprosesser

Magne Jørgensen har satt opp en sjekkliste for "Software Cost Management". Her presenterer han understående rammeverk for kostnadsstyring av systemutviklingsprosesser.

PREPARATION PHASE	
	<ol style="list-style-type: none"> <li>1. Understand estimation problem</li> <li>2. Agree on decisions and assumptions relevant for estimation</li> <li>3. Collect information relevant for estimation</li> <li>4. Select or design the estimation process</li> </ol>
ESTIMATION PHASE	
	<ol style="list-style-type: none"> <li>5. Estimate most likely cost</li> <li>6. Assess uncertainty of cost estimate</li> <li>7. Review estimation process</li> </ol>
APPLICATION PHASE	
	<ol style="list-style-type: none"> <li>8. Apply estimate in bidding</li> <li>9. Apply estimate in planning and budgeting</li> <li>10. Communicate estimate (including communication of uncertainty, bid, plan and budget)</li> <li>11. Control cost (including re-estimation)</li> </ol>
LEARNING PHASE	
	<ol style="list-style-type: none"> <li>12. Learn from estimation feedback</li> </ol>

Tabell 2.5. Jørgensens rammeverk (Jørgensen 2003)

Det er flere interessante trekk i dette rammeverket. For det første ser vi at fokus i prosessen igjen er utvidet. Fra de mer anvendelseskonsentrerte prosessene lansert på 90-tallet se man her på estimering helt fra forberedelse til oppfølging og læring. Dette favner enda bredere enn Boehms 7-punkts prosess. Det andre som er spesielt interessant relatert til mitt hovedfagsarbeid, er punkt 4, nemlig en egen aktivitet som går på å velge eller designe estimeringsprosessen. Ser vi nærmere på to guidelines Jørgensen gir for denne aktiviteten, understrekes nettopp forhold som blir nærmere belyst i mitt arbeid.

*"4.1. Base the selection/design of the estimation process on the availability of relevant estimation information and the estimation situation. Do not select/design an estimation model and then search for the,possibly unavailable or low quality, required input information."*

og

*"4.2. List the available estimation processes and the criterias for selection/ designing them." (Jørgensen 2003) side 2.*

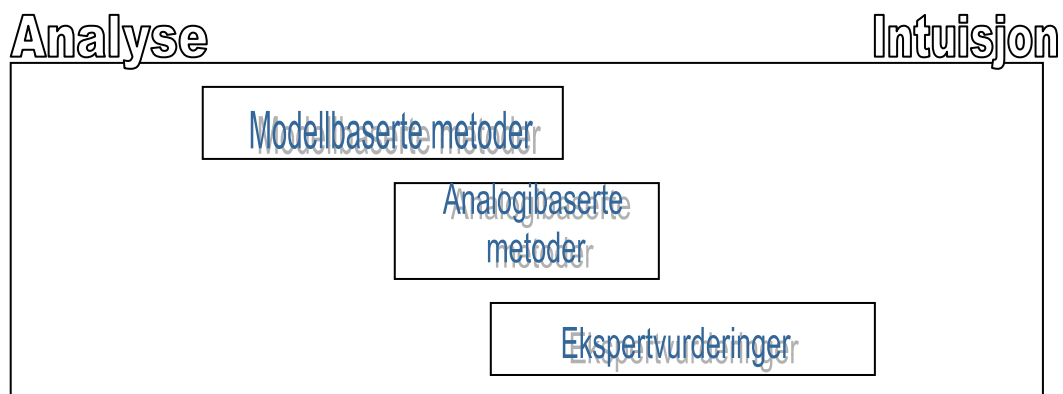
## 2.5. Metodenes forening av analyse og intuisjon

Estimeringsmetodene innen systemutvikling har elementer av både analytisk tilnærming og bruk av intuisjon. Kombinasjonen av at man jobber med prediksjoner, en aktivitet

som i seg selv er meget krevende, og at systemutvikling som fagområde er så komplekst og sammensatt, stiller krav til at metodene må ivareta ulike perspektiver. Hammond belyser i sin bok (Hammond 1996) fordeler ved å søke å forene intuisjon og analyse :

*"I have developed a theory that reduces the rivalry between intuition and analysis and the tension between the coherence and the correspondence theories of truth. That step allow the major approaches to the topic of judgment and decision making to complement, rather than to attempt to displace, each other. " (Hammond 1996) side 228.*

Det er likevel betydelig forskjeller i hvordan metodene søker å forene analyse og intuisjon. Mens de modell-baserte metodene har sitt fotfeste i en analytisk tilnærming, er innslaget av intuisjon utpreget innen ekspertvurderinger. Innen hovedkategoriene finnes også enkeltmetoder med til dels svært ulik tilnærming. Jeg plasserer i modellen under de hovedkategoriene vi har innen systemutvikling i forhold til Hammonds teori om å forene analyse og intuisjon.



Figur 2.3. Estimeringsmetodenes grad av kobling mellom analyse og intuisjon

Ved valg av metode er det nyttig for de ansvarlige å ha et bilde av hvor metodene befinner seg på denne skalaen, da dette forteller noe om metodenes prinsipielle vektlegging. Desto lenger ut til venstre i modellen en metode befinner seg, jo sterkere innslag av formalisme har den. Modellen viser at kategoriseringer av metoder ikke kan sees som en ren hierarkisk inndeling av disjunkte metoder. Modellen viser at jeg ikke vurderer noen av metodene innen systemutvikling til å være ekstreme ytterpunkter. Bak de modellbaserte metodene ligger det alltid et innslag av ekspertvurdering i arbeidet med å sette opp den matematiske formelen for å beregne innsatsbehovet. Et eksempel på dette er COCOMO, der Boehm benyttet seg av ekspertuttalelser i forbindelse med utarbeidelse av modellen (Boehm 1984). Indikatorer for størrelse og produktivitet i de ulike modellbaserte metodene er også basert på eller justert utfra en ekspertvurdering på estimeringstidspunktet.

På den andre siden blir ekspertvurderingene heller ikke praktisert i en ekstremt intuitiv form. Med bakgrunn i litteraturstudiet vil jeg forklare dette med at systemutviklere er sterkt knyttet til tradisjonen innen ingeniørbaserte tilnærming. Ekspertene vil på en mer eller mindre uformell måte, benytte seg av empiri i form av tidligere erfaringer og analysere den informasjon som foreligger om den forestående oppgave.

Når det gjelder Analogibaserte metoder er disse i stor grad tuftet nettopp på ideene om fordelene ved å søke å kombinere styrkene fra hver av de to andre hovedkategoriene (Shepperd, Shofield et al. 1996), og jeg plasserer dem derfor sentralt i modellen. Konklusjon er at man innen fagfeltet har ulike metodekategorier som alle i en viss utstrekning kombinerer analytisk og intuitiv tilnærming.

Jeg mener at det perspektivet denne modellen plasserer metodene i, i altfor liten grad har preget fagfeltet. Det har vært en uheldig polarisering der mange forskere har søkt i retning av å lansere modellbaserte metoder med et sterkest mulig analytisk innslag og betraktet ekspertvurderinger som ren "mage-følelse" basert gjetning. På den andre siden har praktikerne ofte betraktet de modellbaserte metodene som rene akademiske konstruksjoner som krever kvantifiserte indikatorer man i liten utstrekning kan gi.

Hughes (Hughes 1996) påpeker behovet for at metodene bør bygges på en kombinasjon av analytisk, statistisk som ekspertbasert tilnærming. Han mener forskningen bør fokusere mer på hvordan man støtte ekspertene på felter de er svake innen. Briand (Briand, El Emam et al. 1998) har lansert hybrid-modellen COBRA som skal ivareta disse perspektivene. I min tilnærming fokuserer jeg på nytteverdien av å ha en samling av alternative metoder med ulike perspektiver, og heller gjøre et bevisst valg av metode blant de allerede eksisterende alternativene, fremfor at man skal forsøke å utvikle nye ofte komplekse metoder som skal dekke alle hensyn.

### 2.6. Kombinasjon av metoder

I publikasjoner kommenteres potensialet som ligger i å kombinere ulike metoder innen et og samme prosjekt (Boehm 1984; Jørgensen 2003). En måte å kombinere metoder på er å anvende ulike metoder på ulike steg i et prosjekts livssyklus. Motivet for å estimere er ofte vidtfavnende. Man ønsker både et grunnlag å fatte de overordnede beslutningene på, samtidig med at man trenger støtte i den detaljerte ressursplanleggingen. De ulike metodene har styrker og svakheter som gjør at man vanskelig kan tilfredsstille alle krav med bruk av en enkelt metode. Informasjonsgrunnlaget som gradvis vokser seg større og mer detaljert utover i utviklingsforløpet, aktualiserer et eventuelt bytte metode underveis.

En annen måte å kombinere estimeringsmetoder på, er å benytte flere metoder parallelt. Metodene har som jeg vil komme tilbake til komplementære egenskaper, og ser systemutvikling fra ulike perspektiver. For å redusere risikoen for feilestimer er det derfor fornuftig å estimere ved hjelp av flere metoder og la de kvalitetssikre hverandre. Argumentet som i første rekke går mot denne tilnærminsformen, er at ressursbruken til støtteoppgaven estimering vil øke.

Jeg har vært nødt til å begrense meg ved utarbeidelsen av rammeverket, og har derfor ikke som ambisjon å gå dypere inn på dette kompliserte, men meget lovende området. Dette blir for omfattende å dekke dette, men jeg har lagt vekt på å tilrettelegge et rammeverk med en grunnstruktur som tar høyde for dette perspektivet som gjennom litteraturstudiet har utkrystallisert seg som interessant og lovende. Jeg har i rammeverket belyst metodenes styrker og svakheter relatert til konteksten i et prosjekt, og det er fullt mulig å bruke rammeverket til å velge mer enn en enkelt metode.



## 3. STRATEGI FOR VALG AV ESTIMERINGSMETODE

### 3.1. Tilnærminger til å velge metode

Potesialet som ligger i at man kan velge blant et større antall estimeringsmetoder har som gjennomgangen viser i liten grad vært fokusert i forskningen innen fagområdet. Den akademiske tilnærmingen har til nå vært preget av at man har forsøkt å komme frem til estimeringsmetoder som viser seg å være generelt mer treffsikre enn konkurrerende metoder. Sammenligningsstudiene har ensidig fokusert på treffsikkerhet som målekriterium. Følgelig har heller ikke ulike strategier for å foreta et bevisst valg av metode blitt diskutert i særlig grad i litteraturen. Som jeg var inne på i kapittel 2 er det likevel tendenser til at man i nyere publikasjoner begynner å se potensialet som ligger i mer bevissthet rundt hele estimeringsprosessen.

Min angrepsvinkel ved utformingen av et rammeverk for å bygge opp under bevissthet i valg av estimeringsmetode, ble derfor preget av at jeg måtte starte fra grunnen av. Jeg måtte først finne en fornuftig tilnærming til hvordan man kunne legge opp den helhetlige valgprosessen. Et alternativ var å utarbeide en helt ny tilnærming fra bunnen av. Et eksempel på hvordan jeg angrep dette er at jeg brukte ressurser på å vurdere og forsøke å utforme egne Process Patterns (Ambler 1999) for valg av estimeringsmetode innen systemutvikling. Dette ble forsøkt utarbeidet uten at jeg klarte å identifisere interessante mønstre i den etablerte praksis på området. Utarbeidelse av en sjekkliste-basert valgprosess ble også prøvd, og forsøk ble gjort på danne en prosjektpofil-basert valgstrategi.

Det alternativet jeg endte opp med var å basere meg på forskning innen prediksjon fra andre fagområder. Studiene innen dette feltet viste at det er gjort mye innen strategier for valg blant alternative prediksjonsmetoder (Chambers, Mullick et al. 1971; Makridakis, Wheelwright et al. 1983; Mahmoud, DeRoeck et al. 1992; Winklhofer, Diamantopoulos et al. 1996). Armstrongs bok "Principles of Forecasting" (Armstrong 2001) ble sentral i arbeidet, og jeg valgte å ta utgangspunkt i artikkelen "Selecting forecasting methods" (Armstrong 2001), der han gjennomgår seks ulike strategier for valg av prediksjonsmetode. Under følger en vurdering av potensialet i hver av de seks tilnærmingene, etterfulgt av en konklusjon på hvilken av disse jeg fant mest hensiktsmessig å anvende i rammeverket.

### 3.2. Gjennomgang av seks strategier for metodevalg

De seks strategiene som Armstrong presenterer til er :

- Convenience (Overbevisning)
- Market Popularity (Markedspopularitet)
- Structured Judgement (Strukturert Vurdering)
- Statistical Criteria (Statistiske Kriterier)
- Relative Track Records (Resultat på gjennomførte prosjekt)
- Principles from Published Research (Prinsipper fra publisert forskning)

#### 3.2.1. Overbevisning

Det finnes undersøkelser som ser på hvilke metoder de som sitter med ansvar for estimering velger å benytte innen systemutviklingsprosjekter (Heemstra 1992; Kitchenham, Pfleeger et al. 2002), men jeg har ikke funnet nærmere analyser av hvorfor prosjektledere og utviklere velger som de gjør. Siden det ikke finnes generelle etablerte retningslinjer og rammeverk for valg av metode, har utøverne lite å støtte seg til. Enkelte virksomheter har en etablerte politikk som er styrende for valg av metode, men i de fleste sammenhenger vil det være individuelle preferanser ofte i form av prosjektlederens faglige bakgrunn og tidligere erfaringer som bevisst eller ubevisst utgjør grunnlaget for valg av estimeringsmetode.

I forhold til Armstrongs tilnærminger vil jeg plassere en slik tilnærming som valg ut fra overbevisning. Armstrong vurderer dette til å være en billig, men risikable strategi å benytte når man skal velge metode. Dette kan føre til anvendelse av den metode prosjektlederen føler seg mest kompetent innen, men som i liten grad er egnet til de karakteristika som særpreger prosjektet. Anvendelse av overbevisning som valgstrategi kan gi en tendens mot å velge metoder som er enkle og lite ressurskrevende. I omgivelser preget av sterkt tidspress og knappe ressurser, noe man ofte finner innen systemutvikling, er det mest naturlig å velge ofte enkleste vei.

Dette er en hensiktsmessig strategi i tilfeller der man vurderer det som feil å bruke tid til å velge metode, men risikoen forbundet med strategien gjør den lite egnet som basis i et rammeverk der målet nettopp er å oppnå et bevisst og hensiktsmessig valg av metode som ivaretar konteksten prosjektet skal gjennomføres innen.

#### 3.2.2. Markedspopularitet

Prinsippet i denne strategien er å basere valget av metode på at den beste metoden er den som over tid har oppnådd den høyeste popularitet blant aktørene som tidligere har gjort tilsvarende valg. Her har vi en spesiell situasjon innen systemutvikling, der det er stor forskjell i det som har vært populært og preget utviklingen innen forskning, i forhold til det som har oppnådd størst aksept ute blant de som i praksis utøver estimering.

Til tross for en meget sterk fokus på modellbaserte metoder i forskningen og inkorporering av modellbaserte metoder i ulike estimeringsverktøy, har ingen av disse metode oppnådd noen betydelig markedsandeler eller popularitet blant utøverne. Den sterke profileringen av COCOMO i lærebøker og publisert forskning har ikke resultert i en tilsvarende dominans av denne estimeringsmåten i praksis.

Studier viser at (Heemstra 1992; Kitchenham, Pfleeger et al. 2002) ekspertvurderinger er den mest utbredte form for estimering, men måten å gjennomføre ekspertvurderinger er som jeg vil komme tilbake til lite ensartet. Det er her ikke snakk om en homogen måte å estimere på som organisasjoner velger å benytte fordi andre har hatt suksess med den.

Min tilnærming til valg av metode, der jeg understreker nødvendigheten av å velge metode ut fra de faktiske prosjektkarakteristika, taler mot å anvende denne tilnærmingen ved metodevalg. *Markedspopularitet* som valgstrategi blir det stikk motsatte prinsipp, der man skal basere valget på at det beste for deg vil være det som over tid har oppnådd den høyeste populariteten. Armstrong advarer også mot å benytte dette som valgstrategi.

#### 3.2.3. Strukturert Vurdering

Armstrong presenterer *Strukturert Vurdering* som en egnet tilnærming til metodevalg i situasjoner der man har et antall aktuelle metoder å velge blant, og det er mulig å sette opp en liste av kriterier som grunnlag for vurdering av metodene.

I forhold til situasjonen innen systemutviklingsfeltet er *Strukturert Vurdering* en meget interessant strategi. Med en lav bevissthet blant de som er ansvarlige for å velge metode rundt det spektret av alternative estimeringsmetoder som finnes innen systemutvikling, er det en stor fordel med en valgstrategi som eksplisitt viser alternativene man har å velge blant. Når vi i tillegg har hatt det jeg oppfatter som en overdreven fokus mot treffsikkerhet som eneste valgkriterium for valg av metode, vil det være meget gunstig med en valgstrategi som bygger opp under at man skal trekke inn en rekke ulike kriterier som bør påvirke metodevalget.

#### 3.2.4. Statistiske Kriterier

Mye av det publiserte forskningsmaterialet som sammenligner ulike metoder gjør dette med basis i statistiske kriterier. Forskerne vurderer metodene opp mot hverandre ut fra statistiske resultater der treffsikkerheten på estimatet er kvalitetsmålet og ulike datasett av historisk gjennomførte systemutviklingsprosjekter er objektet som estimeringsmetodene anvendes på. Med bruk av ulike målekriterier som for eksempel de godt innarbeidede PRED og MMRE, finner man metodenes treffsikkerhet (Conte, Dunsmore et al. 1986; Shepperd, Shofield et al. 1996).

Et spørsmål man bør stille seg i forhold til deler av disse undersøkelsene er hvorvidt man i ettertid klarer å simulere den situasjonen man reelt har på det tidspunkt estimatene gjøres. Tallene på det faktiske forbruket i prosjektet er også en mulig feilkilde ved slik bruk av empiriske data. Rapportering av forbruk av innsatsen er gjenstand for svært mange påvirkninger og usikkerhetsmomenter i tilknytning til systemutviklingsprosjekter, at det kan stilles spørsmål om dette kan gi skjevheter i forhold til vurderingen av treffsikkerheten til metodene. Hovedproblemet er likevel at resultatene i sammenligningsstudier spriker, selv om treffsikkerhet benyttes som mål i de fleste studiene.

#### 3.2.5. Resultat på gjennomførte prosjekter

Dette er en valgstrategi der man sammenligner ytelsen til de ulike metodene med anvendelsen av systematiske, balanserte og pålitelige prosedyrer. Armstrong understreker at det ikke er en vurdering av hvor fornøyd man i en organisasjon over tid har vært med anvendelsen av enkeltmetoder. I stedet er det en vitenskapelig sammenligning av metodenes ytelse på et nærmere angitt målekriterium ut fra et sett av organisasjonens egne historiske data. Utgangspunktet er at disse historiske dataene skal være representative som en generalisering i forhold til fremtiden.

Mangel på kompetanse til å definere prosedyrer for å evaluere metodene og mangel på historiske data av tiltrekkelig kvalitet på gjennomførte prosjekter taler mot denne strategien. Den sterke utviklingen innen fagområdet sår etter min vurdering tvil om de historiske dataenes representativitet i forhold til kommende prosjekter. Jeg vurderer derfor denne valgstrategien til ikke å være blant de mest egnede for den gjennomsnittlige virksomhet som driver systemutvikling.

#### 3.2.6. Prinsipper fra publisert forskning

Å basere et metodevalg på en strategi der man anvender de prinsipper som den akkumulerte forskningen innen fagfeltet har gitt, virker som en lovende tilnærming. Armstrong har selv anvendt denne valgstrategien for å komme frem til et forslag til seleksjonstre for valg mellom prediksjonsmetoder. For å kunne anvende dette fordres at man innen forskningen har en tilstrekkelig mengde av publikasjoner som har gjort så vel brede som dype studier innen dette spesifikke fagfeltet. Jeg har ikke funnet et tilstrekkelig antall publikasjoner med en sterk nok fokusering på valg av estimeringsmetode innen systemutvikling til å kunne anvende denne valgstrategien, selv om den etter Armstrongs vurdering er meget lovende.

#### 3.3. Bruk av *Strukturert Vurdering* i rammeverket

Ut fra min gjennomgang av de seks ulike strategiene, ser det for meg ut til at man innen vårt fagfelt har en tendens mot at praktikerne benytter seg av overbevisning ved valg av estimeringsmetode, mens forskere i stor grad har tendert mot å forsøke å bygge opp en objektiv metodevurdering rundt statistiske kriterier. Her har treffsikkerhet vært det dominerende vurderingskriteriet. Ingen av disse tilnærmingene virker til å ha ført til noen utbredt bevisst holdning til valg av estimeringsmetode. Armstrong gir heller ingen anbefaling i retning av noen av disse valgstrategiene.

Markedspopularitet er for meg en helt uaktuell valgstrategi, ut fra min vurdering av at den aktuelle prosjekt og organisasjonskontekst i stor grad bør få prege valget av metode. Relative Track Record er en strategi som er interessant ut fra at man bygger et metodevalg på de resultater enkeltmetoder kan vise til innen den gitte organisasjonskonteksten. Et spørsmål man kan stille seg er likevel om en så sterk kobling mot historiske data er gunstig innen vårt fagområde som er preget av stor teknologisk utvikling og stadige skifter i anvendelsesområder. Strategien vil også være krevende å anvende i praksis, da virksomhetene måtte besitte kompetanse til å forme en god vurderingsprosedyre, og skaffe nok informasjonsgrunnlag til å gjennomføre strategien på en vitenskapelig måte.

Følgelig gjenstår de to strategiene som av Armstrong begge betraktes som lovende – *Strukturert Vurdering* og Prinsipper fra publisert forskning. Jeg har valgt å benytte *Strukturert Vurdering* da denne åpner for et sterkt innslag kontekstuell bedømmelse ved at man velger ut kriteriene selv, samt at den fremtvinger en bevisstgjøring av mangfoldet i de alternative metodene vi har. Som jeg var inne på over har forskningen vært for homogen til å utgjøre et godt grunnlag for utlede prinsipper ut fra.

I forhold til Armstrongs beskrivelser av *Strukturert Vurdering*, har jeg i det videre arbeidet forsøkt å tilpasse denne tilnærmingen til metodevalg til de faktiske forhold man opplever i valgøyeblikket innen systemutviklingsprosjekter. Videre er kriteriene endret fra å ha generell fokus, til å rette seg spesielt mot å ivareta organisasjons- og prosjektkontekst

Briand og Wieczorek har brukt en lignende tilnærming i sitt arbeid (Briand and Wieczorek 2002). Deres rammeverk for sammenligning og evaluering av estimeringsmetoder innen systemutvikling, har 16 evalueringskriterier som anvendes i vurderingen av 14 alternative metoder. Deres tilnærming preges likevel av at det gis en

### 3. STRATEGI FOR VALG AV ESTIMERINGSMETODE

---

en mer generell evaluering av metodene. Jeg lanserer en alternativt sett av vurderingskriterier for valg av metode. Disse er orientert mot å se på de kontekstuelle forholdene. Når det gjelder de alternative metodene man har å velge blant er forskjellen stor. Jeg opererer med langt bredere inndeling i kategorier, mens de har klar skjevhet mot modellbaserte metoder.

## 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

I følge Armstrong (Armstrong 2001) er kjennskap til de aktuelle prediksjonsmetodene nødvendig for å kunne bruke *Strukturert Vurdering* som valgstrategi. For å kunne anvende denne strategien til å gjøre et bevisst valg av estimeringsmetode innen systemutvikling, er det derfor en forutsetning at det finns en oversikt over alternativene man kan velge blant. Ut fra det jeg har funnet i litteraturstudiet eksisterer ingen kategorisering av metodene for estimering innen systemutvikling som godt nok viser bredden i de metodene som reelt er tilgjengelig. Den kategoriseringen jeg i dette kapitlet foreslår har til hensikt å gi en overordnet oversikt over spekteret av metoder som er relevante innen estimering av systemutviklingsprosjekter.

- |  |
|--|
| <ol style="list-style-type: none"><li>1. Algorithmic Models</li><li>2. Expert Judgement</li><li>3. Analogy</li><li>4. Parkinson</li><li>5. Price to Win</li><li>6. Top-Down</li><li>7. Bottom-Up</li></ol> |
|--|

*Tabell 4.1. Barry Boehms syv klasser*

Det eksisterer ingen objektivt og etablert klassifisering av estimeringsmetodene for systemutvikling. Etter at Barry Boehm (Boehm 1981) introduserte syv ulike klasser av estimeringsteknikker, har ulike alternative kategoriseringer blitt foreslått og anvendt innen fagfeltet (Fairley 1992; Heemstra 1992; Fenton and Pfleeger 1996; Walkerden and Jeffery 1997; Briand and Wiczorek 2002). Flere av disse har med utgangspunkt i Boehms tidlige klassifisering forsøkt å forbedre kategoriseringen, men etter min vurdering er altså ingen av dem dekkende som grunnlag for å foreta metodevalg basert på *Strukturert Vurdering*.

Felles for de mest detaljerte av kategoriseringene (Walkerden and Jeffery 1997; Briand and Wiczorek 2002) er at det gis en svært skjev fremstilling, med en overdreven tyngde på modellbaserte metoder. Det gis en detaljert inndeling innen modellbaserte metoder, mens spesielt ekspertvurderinger betraktes som en samlet mengde av diverse individuelle og uklassifiserte tilnærminger uten noen form for nærmere inndeling. Walkerden og Jeffery velger sågar å se bort fra denne estimeringsformen i sin oversikt, til tross for at de erkjenner at dette er den i praksis mest utbredte tilnærmingen.

*"Expert judgment is also recognized as a prediction method (e.g., Boehm, 1981; Heemstra, 1992). Experts may employ one or more of the other methods in making predictions, either informally or formally. It is likely that expert judgment is employed to make predictions whenever an expert is available. Expert judgment is not included in the framework for selecting prediction methods, as this method cannot easily be characterized, and it is assumed that it is selected whenever experts are available."*  
(Walkerden and Jeffery 1997) side 71-72.

I mitt kategoriseringsskjema ønsker jeg nettopp å vektlegge den totale bredden i tilgjengelige estimeringsmetoder. Siden ekspertvurderinger er så utbredt i praksis, er det spesielt interessant å kartlegge de ulike tilnærminger ekspertene kan bruke når de estimerer. For å få frem det mangfold av ulike tilnærminger man har innen ekspertvurderinger har jeg her derfor lansert et nytt forslag til inndeling av disse, mens jeg for modellbaserte metoder har vurdert de eksisterende inndelinger og basert inndelingen på den mest egnede. Mitt forslag til kategorisering er ikke ment å ha generell gyldighet og anvendelighet innen estimering. Det er laget med henblikk på å inngå i mitt rammeverk for valg av metode.

For å gi et lettfattelig oversiktsbilde er kategoriseringen satt opp med utgangspunkt i en hierarkisk inndeling. Her skisseres de sentrale alternative kategorier av metoder innen fagfeltet. Målet er å øke bevisstheten rundt bredden av alternative metoder, og hierarkiet må ikke tolkes som noen steng formell inndeling i disjunkte metodekategorier, der hver enkelt metode entydig skal kunne plasseres i en enkelt sub-kategori. Relatert til *Strukturerte Vurdering* skal denne inndelingen sikre at de estimeringsansvarlige ikke bare forholder seg til metoder de allerede har kjennskap til når de velger metode.

Videre i dette kapitlet presenterer jeg først de tre hovedkategoriene jeg har delt metodene inn i, og henviser til funn i min litteraturstudie som utdyper den status hver av disse har innen fagfeltet med fokus på forhold som angår valg av metode. Deretter foreslår jeg en sub-kategorisering og gir hver av disse en beskrivelse og kommenterer sterke og svake sider ved denne typen metoder som er påpekt i publisert forskning. Dette vil nemlig stå helt sentralt ved et metodevalg, der karakteristika ved den gitte kontekst skal kobles nettopp til metodenes sterke og svake sider for å finne den mest egnede metoden. Innen ekspertvurderinger har jeg kommet til at den rent hierarkiske struktur ikke godt nok dekker mangfoldet, så her presenteres i tillegg noen sentrale karakteristika som supplerer den hierarkiske inndelingen. Dette er med på å synliggjøre det mangfoldet av alternative tilnærminger innen ekspertvurderinger som hittil har vært mangelfullt belyst.

### 4.1. Hovedkategorier

Jeg har vurdert og bearbeidet de publiserte forslagene til kategorisering jeg nevnte innledningsvis i kapitlet, og her funnet det mest hensiktsmessig å dele estimeringsmetodene inn i tre hovedkategorier:

- Modellbaserte metoder
- Analogibaserte metoder
- Ekspertvurderinger

Disse utgjør i valgsammenheng viktige hovedretninger. Det understrekes på ny at dette ikke er disjunkte kategorier, men av hensyn til å kunne foreta et kontekstavhengig valg er det likevel viktig å gi en lettforståelige inndeling som fanger de viktigste skillene blant de eksisterende metodene. Skillet mellom Modellbasert metoder og Ekspertvurderinger er veletablert i så vel lærebøker som publisert forskning. De representerer metoder med ulike perspektiver på estimering, nemlig en matematisk og

## 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

---

formalistisk kontra en kognitiv og humanistisk tilnærming. Selv om de eksisterende estimeringsmetodene innen systemutvikling som nevnt ikke representerer rendyrkede former av dette, er det viktig at de som velger metode forstår de ulike perspektivene metodene representerer.

At analogibaserte metoder er skilt ut som en egen hovedkategori er ikke like åpenbart som skillet mellom modell og ekspert. Det kan argumenteres for at også analogibasert estimering bygger på anvendelse av modeller og dermed hører til under modellbaserte metoder (Briand and Wiczorek 2002). Andre ser analogibruk i en mer uformell sammenheng, og betrakter analogier som erfaringer eksperter på en ustrukturert måte benytter seg av når kommer frem til sine estimater (Hughes 1996). Analogibasert estimering kan dermed på mange måter betraktes som en mellomting mellom modellbasert estimering og ekspertvurderinger.

I mitt arbeid velger jeg likevel å skille ut analogibaserte metoder som en egen kategori metoder som forsøker å forene styrkene fra de to andre kategoriene, men som likevel har sine særpreg. Riktignok er det vanlig å anvende en type modeller også ved analogibasert estimering, men disse brukes til å identifisere det prosjektet i en samling av gjennomførte prosjekter som ut fra estimeringskritiske forhold er mest sammenfallende med utviklingsprosjektet. Dette skiller seg fra det jeg legger i Modellbaserte estimeringsmetoder, der den matematiske modellen beregner selve innsatsestimatet for den forestående oppgaven.

Modellene som brukes til å finne det mest sammenfallende prosjektet er ofte basert på målinger av euklidske distanse (Shepperd and Schofield 1997). Det eller de prosjekter med den minste euklidske distanse til det nye prosjektet velges ut som analogiprojekt. Estimater baserer seg i neste omgang på en justert versjon av det reelle historisk rapporterte forbruket fra denne analogien. Modellen brukes her altså ikke for å beregne selve estimatet, men snarere til å identifisere det mest sammenfallende prosjektet. Denne forskjellen påvirker metodevalget da det stilles ulike krav til omfanget av og grad av struktur på de tilgjengelige historiske data.

Alternativt kan man velge å identifisere analogiprojekt(er) ved å få en ekspert til å peke ut det historiske prosjekt som egner seg best til å brukes som direkte analogi. Slik anvendelse av eksperter skiller seg også fra det jeg her kategoriserer som ekspertvurderinger, da eksperten innen analogibaserte metoder bruker sin kompetanse til å gjøre en seleksjon av analogiprojekt, fremfor å gi sine estimater på innsatsbehov. Det som særpreger analogibasert estimering blir således at selve estimatet i større grad utledes av historiske data på et gjennomført prosjekt, mens det er identifiseringen av analogiprojektet som blir det sentrale. Den tradisjonelle form for uformell analogibruk plasserer jeg i min inndeling inn under ekspertvurderinger.

### 4.2. Status for hver av de tre hovedkategoriene

#### 4.2.1. Modellbasert estimering

Av de tre hovedkategoriene har de modellbaserte metodene definitivt hatt den dominerende rollen innen forskningen (Hughes 1996; Gray, MacDonnell et al. 1999). Et stort antall metoder med tilhørende algoritmiske modeller og estimeringsteknikker er lansert og videreutviklet av ulike forskningsmiljøer. Felles for modellene er at de



#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

---

baserer seg på at det eksisterer en matematisk sammenheng mellom innsatsen som trengs til å utvikle programvare og en eller flere indikatorer. For å oppnå et treffsikkert estimat må indikatoren(e) kunne måles eller forutsies med stor nøyaktighet på estimeringstidspunktet. En hovedutfordring innen modellbaserte metoder ligger i få den nødvendige kvalitet og treffsikkerhet på disse indikatorene, og dette blir derfor et vesentlig forhold i vurderingen av hvorvidt disse metodene er relevant i en gitt prosjektkontekst. Kritikken mot modeller som kombinerer prediksjoner av størrelse og produktivitet går på at man vanskelig kan oppnå et godt estimat når modellen man benytter også baserer seg på usikre prediksjoner (Walkerden and Jeffery 1997).

Størrelsen på programvaren som skal utvikles er den mest utbredte indikatoren som anvendes i modellene for estimering av systemutviklingsprosjekter, da det har vært allment akseptert innen fagfeltet at det er klar sammenheng mellom størrelsen på en programvare og den innsats som kreves for å utvikle den. Hvorvidt størrelse virkelig har så sterke føringer for innsats er for øvrig et forhold som enkelte nå har begynt å sette spørsmålsteget ved (Dolado 2001). Følgende sitat er likevel representativt for store deler av publikasjonene innen området og viser utfordringene man her sliter med :

*"From a general perspective, it is clear that sizing is one of the key issues to address to devise accurate, cost-effective ways of predicting resource expenditures. Though many pertinent ideas have been reported in the literature, no existing solution can claim to provide the universal answer."*  
(Briand and Wieczorek 2002), side 23.

Problemet ligger i at prediksjonen rundt størrelsen i seg selv er svært usikker på det tidspunktet man trenger estimatene. Antall kodelinjer (LOC) i den ferdige programvareløsningen og angivelse av funksjonspoenger (FP) knyttet til den spesifisering som skal tilfredsstilles har vært utbredte størrelsesenheter i modellbaserte metoder. Nyere forskning ser på tilpasning av modellene til moderne analysemetoder. Eksempelvis ha man gjort lovende forsøk med å anvende Use-Case som utgangspunkt for prediksjoner av størrelse og i neste omgang innsatsestimat (Anda B, Angelvik E et al. 2002).

Ved siden av størrelse er produktivitet en sentral indikator i mange modellbaserte metoder. Mange modeller forsøker å inkorporere produktiviteten i den aktuelle organisasjonen i form av at man legger inn et sett lokalt kalibrerte kostnadsdrivere. Kompleksiteten rundt måling av produktiviteten i systemutviklingsprosjekter er stor, og de eksisterende estimeringsmodellene sviker med hensyn på å reflektere de faktorer som påvirker produktiviteten (Maxwell, van Wassenhove et al. 1996).

##### 4.2.2. Analogibasert estimering

Filosofien bak denne måten å estimere er at historien gjentar seg. Walkerden angir fire faktorer som samlet er avgjørende for treffsikkerheten man kan oppnå ved å benytte analogibaserte metoder, og som derfor må trekkes inn ved metodevalget.

- Tilgjengelighet på passende analogier
- Egnetheten til strategien for å velge ut analogien
- Håndteringen av forskjellene mellom analogi og estimeringsobjekt under utledning av estimatet
- Nøyaktigheten på dataene som anvendes både som analogi og estimeringsobjekt (Walkerden and Jeffery 1999).

#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

---

I forhold til valg av metode blir tilgjengeligheten på historiske data et helt sentralt punkt i forhold til om man vurderer analigbaserte metoder som relevante. Så tidlig som i 1981 skisserte Boehm (Boehm 1981) bruk av analogier som en egen tilnærming til estimering. Følgende fremtidsvisjoner presentert av Fairleys viste seg senere å slå til i form av økning i forskningsfokus på området :

*"Increased collection, analysis, and retention of historical data, resulting in more accurate rules of thumb and analogies, and calibration of estimation techniques to local conditions."* og

*"More sophisticated tools to support analogy based estimation; for example rules to select analogous projects from data repositories; better measures of goodness for analogies."* (Fairley 1992) side 390.

Det ble på 1990-tallet utformet metoder der en formalisert og direkte anvendelse av historiske data fra analogiprojekter ble tatt i bruk (Shepperd, Shofield et al. 1996). Hittil har disse estimeringsmetodene i første rekke blitt anvendt av virksomheter som etablerer en prosjektdatabase med relevante interne prosjekter. En utvikling i retning av prosjektdatabaser på bransjenivå kunne gitt metodene et større nedslagsfelt, men et hinder for utviklingen her ser ut til å være at sensitivitet på informasjon og usikkerhet på datakvalitet gjør at virksomheter foretrekker å holde informasjonen internt.

##### 4.2.3. Ekspertvurderinger

Jørgensen (Jørgensen 2002) oppsummerer relevant forskning innen ekspertvurderinger. Her fremgår det at det er svært få publikasjoner på ekspertvurderinger i forhold til den samlede mengde publikasjoner innen estimering av systemutviklingsprosjekter. Tilsvarende tendens fant jeg i min gjennomgang av fagbøker. Dette står i kontrast til at oppsummeringen samtidig viser til undersøkelser som alle bekrefter at ekspertestimering er den estimeringsmetode som blir foretrukket i praksis. For ekspertvurderinger er altså status at denne er svært utbredt i praksis, men forskningen har vært meget begrenset

Innen andre fagfelter nådde ekspertvurderinger allerede på midten av 80-tallet en aksept som vi først de siste årene har sett tendenser til innen systemutvikling. Da ble ekspertise og dens rolle i tilknytning til prediksjoner gjenstand for seriøs forskning (Ferrell 1985).

Ekspertvurderinger her beskrivet som:

*"Expert judgement is data given by an expert in response to a technical problem. An expert is a person who has background in the subject area and is recognized by his or her peers or those conducting the study as qualified to answer questions."* (Meyer and Booker 1991) side 3.

I samme bok vises det til at ekspertvurderinger innen en rekke fagfelt benyttes når andre kilder som målinger, observasjoner, eksperimenter og simuleringer ikke er tilgjengelige, og at ekspertvurderinger egner seg godt innen nye, sjeldne, komplekse eller på andre måter dårlig forståtte fenomener. Meyer mener altså at ekspertvurderinger i første rekke kommer til sin rett der øvrige prediksjonsmodeller ut fra kontekstuelle forhold ikke er anvendbare. Jørgensen (Jørgensen 2003) nyanserer dette bildet med å vise til at ekspertvurderinger innen vårt fagfelt ikke er systematisk svakere enn modellbaserte metoder selv i situasjoner der man er i stand til å gi

#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

modellene nøyaktig input. Min tilnærming er at en nyansert vurdering av konteksten er nødvendig for å finne ut av om, og eventuelt i hvilken form, ekspertvurderinger bør benyttes i et prosjekt.

Ekspertvurderingene er i mange sammenhenger knyttet til at enkeltindivider gir estimer basert individets kompetanse og erfaring. Den tilsynelatende usystematiske og individuelle måten å foreta estimering på står i sterk kontrast til den fokus forskerne innen vårt område har hatt i retning av å søke etter analytiske, generelle og repeterbare estimeringsprosesser som kan forklares overfor omgivelsene. Forskerne ser ut til ha oversett at de individuelle ekspertestimatene er bare utgjør en mulig tilnærming til ekspertestimering innen vårt fagfelt.

Boehm beskrev riktignok tidlig Delphi prosessen som et alternativ til individuelle ekspertestimer. Han foreslo Wide-band Delphi Technique som en aktuell måte å tilpasse gruppeestimeringen til de forhold man har innen systemutvikling(Boehm 1981). Moløkken (Moløkken and Jørgensen 2003) kommenterer at denne formen er blitt hyppig referert uten at den er blitt gjenstand for praktisk anvendelse eller ytterligere forskning.

Enkelte publikasjoner har kommentert styrkene ved og omtalt utbredelsen av ekspertvurderinger, sett på hvem som benyttes som eksperter og til dels sett på hvordan ekspertene jobber når de estimerer (Lederer, Mirani et al. 1990; Vicinanza, Mukhopadhyay et al. 1991; Heemstra 1992; Lederer and Prasad 1992; Hughes 1996; Lederer and Prasad 1998; Gray, MacDonnell et al. 1999). En survey rundt eksperters måte å estimere på foretatt av Hughes innen en organisasjon viste følgende rangering for hvor viktig ulike informasjonskilder var for de individuelle ekspertene ved utarbeidelse av estimatene :

Information used	Summed rating
1. Design requirements	84
2. Resources available	30
3. Base Product/source code	28
4. Software tools available	16
5. Previous history of the product (e.g. previous changes)	16
6. Size/complexity of the new function	9
7. Similar previous implementations	8
8. Documentation impacted	7
9. Amount of midified v. new code	6
10. Test site availability	5
11. Bug reports in the software to be amended	4
12. Deadline pressure	4
13. Customer for the product	4
14. Re-usability requirements	3
15. Expected life-time of the product	3
16. Quality requirements	2

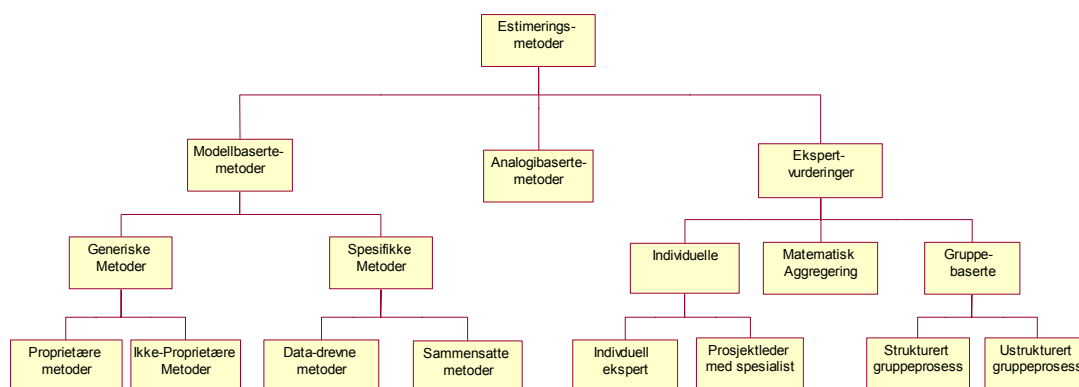
*Tabell 4.2. Rangering av informasjonskilder brukt ved ekspertvurdering (Hughes 1996)*

Det er likevel først i arbeider som (Höst and Wohlin 1997; Höst and Wohlin 1998; Jørgensen and Sjøberg 2001; Jørgensen 2002; Jørgensen, Teigen et al. 2002; Moløkken 2002; Moløkken and Jørgensen 2002; Jørgensen 2003) at man går dypere inn i denne estimeringsformen og ser på alternative måter å gjennomføre ekspertestimer på innen systemutvikling.

### 4.3. Subkategorisering av metodene

I likhet med modellbaserte metoder, kan verken analogibaserte metoder eller ekspertvurderinger vise til et generelt tilfredsstillende nivå av treffsikkerhet på estimatene. Status er dermed at det finnes tre hovedkategorier metoder der ingen gir generelt god treffsikkerhet, og hvor det innbyrdes styrkeforholdet mellom de tre kategoriene varierer i ulike publikasjoner (Jørgensen 2002). Delingen i de tre hovedkategoriene utgjør en god grunnstruktur i sammenheng med valg av metode, men for å støtte valgprosessen i en *Strukturert Vurdering* trengs en finere inndeling.

Dette er mitt forslag til kategorisering som videre i kapitlet vil bli gjennomgått i detalj :



Figur 4.1. Kategorisering av estimeringsmetodene

### 4.4. Modellbaserte metoder

Jeg kommenterte innledningsvis i kapitlet at ingen eksisterende kategoriseringer godt nok viser bredden i tilgjengelige metoder for estimering av systemutviklingsprosjekter. Innen hovedkategorien Modellbaserte metoder derimot, eksisterer det en rekke forslag til aktuelle kategoriseringer som i tilstrekkelig grad viser mangfoldet innen denne typen metoder. Modellbasert estimering har vært sterkt fokusert i forskningen og det finnes her en rekke publikasjoner som foreslår spesifikke kategoriseringer. Mitt arbeid har her gått ut på å vurdere de ulike forslagene, og finne frem til den eksisterende kategoriseringen som best tilfredsstiller mitt formål. Fem ulike kategoriseringer fra ulike tidsepoker ble nærmere vurdert før valget ble gjort.

#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

---

Boehm (Boehm 1984) delte kategorien algoritmiske modeller inn på følgende måte :

- Lineære modeller
- Multiplikative modeller
- Analytiske modeller
- Tabulariske modeller
- Sammensatte modeller

Kemerer (Kemerer 1987) tar også for seg algoritmiske modeller og skiller mellom de modeller som har antall linjekoder (SLOC) som primærinput, og de som ikke har det. Videre skilles det mellom proprietære og ikke proprietære modeller. Et annet eksempel er Fairley (Fairley 1992) som presenterer en inndeling av det han omtaler som estimeringsteknikker, og kategoriserer disse i empiriske teknikker, regresjonsbaserte teknikker og teoribaserte teknikker. Kategoriseringen favner det jeg her omtaler som metoder.

Walkerden og Jeffery (Walkerden and Jeffery 1997) har kategorisert det jeg omtaler som modellbaserte metoder inn i fire :

- Empiriske parametriske modeller
- Empiriske ikke-parametriske modeller
- Teoretiske modeller
- Heuristikker.

Den siste inndelingen jeg har vurdert er Briand og Wieczorek (Briand and Wieczorek 2002) som deler modellbaserte metoder i generiske og spesifikke metoder. De generiske metodene skilles som hos Kemerer i proprietære og ikke proprietære metoder, mens de spesifikke metodene skilles i data-drevne metoder og sammensatte metoder.

Disse fem kategoriseringene er satt opp gjennom et tidsvindu på 18 år. Det er gjort ulike prioriteringer med hensyn til hva som danner grunnlaget for inndelingen. Ser vi for eksempel på COCOMO, plasserer Boehm denne som en sammensatt modell, Kemerer omtaler dette som en linjekodebasert ikke-proprietær modell. Hos Fairley går den som en regresjonsteknikk, mens Walkerden plasserer den under empiriske parametriske modeller. Briand kategoriserer COCOMO som en ikke-proprietær generisk modell. Terminologibruken og kategoriseringsprinsippene over kan stå som et eksempel på at det for de som utøver estimering i praksis er problematisk å hente ut anvendbare prinsipper fra publisert forskning. I forhold til å velge metode blir man overfløymet med ulike inndelinger.

Relatert til metodevalg mener jeg den inndelingen som Briand og Wieczorek gjør er best egnet (Briand and Wieczorek 2002). Inndeling er den nyeste, og tar derfor hensyn til "state-of the art" innen fagfeltet. Det er ingen krav til at utøverne har noen statistisk eller matematisk kunnskap for å kunne forstå inndelingen. Jeg vil også trekke frem at det i valgsammenheng er viktig å ivareta skillet mellom generiske og proprietære modellene, da muligheten for innsyn i og forståelse for hvordan estimatene fremkommer har påvirkning på metodevalget. Det siste argumentet for å velge denne sub-kategoriseringen, er at Briand og Wieczorek i sitt arbeid har gjort grundig evaluering av ulike metoder innen hver av kategoriene, noe jeg nyttegjør meg i mitt rammeverk. Jeg har på denne bakgrunn valgt å bruke deres sub-kategorisering.

I forhold til figur 2.1. side 16 som viser hele klassifiseringen til Briand og Wiecek, er det den venstre grenen som starter med "Modell-Based Methods" som jeg benytter meg av. Denne er oversatt og brukt i min kategorisering i figur 4.1. side 36 under Modellbaserte metoder. **I gjennomgangen av den enkelte sub-kategori av disse metodene vil omtalen av styrker og svakheter ved den enkelte modellbaserte metode være basert på Briand og Wiecek sin evaluering. Unntaket er de tilfeller der jeg presiserer at det er mine egne vurderinger eller der jeg konkret angir en annen kilde.**

### 4.4.1. Generelle svakheter og styrker ved de modellbaserte metodene

#### Svakheter relatert til et kontekstavhengig metodevalg :

De modellbaserte metodene er som tidligere vist meget sensitive i forhold til nøyaktighetsnivået på den informasjonen som gis. Dette gjelder så vel detaljeringsnivå som presisjonen på inputen. Mens man ved ekspertvurderinger kan skjønnsvurdere informasjonsgrunnlaget, og ta høyde for den usikkerhet man ser ligger i dette, vil de modellbaserte metodene direkte rammes av feil input og produsere estimater som ikke er representative (Fairley 2002). Jeg vurderer dette generell svakhet som i mange tilfeller vil diskvalifisere alle de modellbaserte metodene. Det har vært riktignok vært en utvikling mot å gjøre metodene mer fleksible med hensyn til hvilken input modellene krever.

*"Inaccurate sizes will translate directly into inaccurate cost and schedule estimates even if all software models behaved perfectly. Fortunately there is a variety of different size estimation techniques available, and some have been shown to be accurate, at least for limited applications." (Ferens 1999) side 27.*

Avhengighetsforholdet til kvaliteten på input er likevel så stor at jeg mener metodene kan sies å være uegnede i de tilfeller der det råder stor mangel på eller hefter stor usikkerhet ved den input man har mulighet for å fremskaffe på estimeringstidspunktet i en gitt kontekst

#### Styrker relatert til et kontekstavhengig metodevalg :

Uavhengig av hvilken modellbasert metode som velges vil man i stor grad oppnå en automatisert og effektiv estimeringsprosess. Metodene stiller ulike krav til forberedende aktiviteter, men selve estimeringen er i stor grad automatiseres. Det eneste man må gjøre er å fremskaffe den relevante input. Deretter vil metoden gi estimater på de ulike detaljnivåer. Gjentakelser av estimeringen og ajourføring av estimater i forhold til at ny informasjon har kommet frem er enkelt med modellbaserte metoder.

### 4.4.2. Proprietære metoder

Dette er metoder som ut fra et klart definert sett av input variable genererer et estimat. Det spesielle med denne kategorien metoder er at selve beregningsmekanismen er hemmeligholdt. Kun de som har laget modellene kjenner til hvordan estimatet kalkuleres. Disse modellene selges ofte innbakt i et dataverktøy. Disse metodene gir ut fra registrerte input-verdier automatisk både innsatsestimat, kostnadsestimat og varighetsestimat. .

**Svakheter relatert til et kontekstavhengig metodevalg:**

1. Ikke innblikk i estimatberegning.

Utøverne må her godta at man ikke har noen tilgang til det som skjer mellom angivelse av en avkrevd input og det endelige estimat. Man må stole på at de matematiske beregningsmekanismene i modellen. Slike estimater vil i liten grad være egnet for justeringer da man vanskelig kan gjøre kvalifiserte endringer på estimater man ikke vet hvordan er beregnet.

**Styrker relatert til et kontekstavhengig metodevalg:**

1. Lite avhengig av estimeringsekspertise,

Innen disse metodene gis det klare spesifikasjoner på den input som forventes. Siden all estimering foregår skjult for brukeren, stiller metodene minimale krav til teoretisk og praktisk estimeringskompetanse.

2. Bred anvendelse

Metodene markedsføres i stor grad til å dekke alle applikasjonstyper og prosjektstørrelser. Det finnes lite støtte for dette argumentet i empiriske undersøkelser.

**4.4.3. Ikke-proprietære metoder**

Innen ikke-proprietære metoder finner vi anvendelse av ulike matematiske modeller som kalkulerer et innsatsestimat på bakgrunn av et sett av indikatorer. Det som skiller disse fra de proprietære metodene er at modellene er publisert og offentlig tilgjengelige. Dette har ført til at de metoder som hører inn under denne kategorien i langt større grad er blitt gjenstand for forskning basert på empiriske undersøkelser. Det er naturlig å trekke frem COCOMO et eksempel her. Den har fått stor utbredelse i lærebøker og i publiserte forskningsartikler.

**Svakheter relatert til et kontekstavhengig metodevalg :**

1. Lite egnet til tidlige estimater og ved iterative utviklingsmodeller.

Modellene har sin opprinnelse knyttet til tradisjonell fossefallsbasert systemutvikling (Boehm and Sullivan 1999) der tidlige avklaringer rundt omfanget av oppgaven stod sterkere enn i de mer moderne iterative prosessene. Tidlige estimater støttes dermed dårlig av disse metodene, selv om metoder som COCOMO er forsøkt videreutviklet for å holde følge med utviklingen. (Boehm, Abts et al. 2000).

2. Avhengig av kalibrering

Det har lenge vært kjent at disse metodene er avhengig av at de kalibreres til den lokale konteksten før de anvendes (Kemerer 1987).

**Styrker relatert til et kontekstavhengig metodevalg :**

1. Lite avhengig av estimeringsekspertise

Modellene automatiserer selve estimeringsprosedyren, og behovet for ekspertise innen det å estimere en oppgave er liten.

2. Ivaretar kompleksitet og størrelse

Ved utformingen av modellene har man lagt vekt på at alle forhold ved et systemutviklingsprosjekt skal inngå. Slike metoder skalerer derfor godt til større prosjekter, og vil kunne ivareta totaliteten på en rimelig god måte også når kompleksiteten blir stor.

### 3. Er dokumenterte

Lederer anbefaler å stole på dokumenterte fakta, standarder og enkle aritmetiske formler ved estimering. Han mener dette gir større trygghet enn man kan oppnå ved gjetning, intuisjon, støtte i individuelle hukommelse og ved mer komplekse formler. (Lederer and Prasad 1992).

#### 4.4.4. Data-drevne metoder

I kategorien data-drevne metoder plasserer Briand og Wieczorek modeller som er utledet fra data. Det er her snakk om å forme en estimeringsmodell ut fra et sett av data. Disse metodene er sterkt fokusert på anvendelse av regresjonsanalyse for å komme frem til modeller som best mulig uttrykker forholdet mellom den avhengige variabelen (estimatet) og de uavhengige variablene (indikatorene). Etter å ha lansert en modell som spesifiserer forholdet mellom variablene, tilpasses koeffisientene i modellen ved bruk av regresjonsteknikker på et sett av kjente og relevante data. Den enkleste formen, Ordinær Least-Square regression (OLS) har vist relativt gode resultater, men for å takle kompleksiteten innen estimering av systemutvikling har forskere med sterk fokus på anvendelse av statistiske metoder lansert et antall ulike data-drevne metoder.

##### Svakheter relatert til et kontekstavhengig metodevalg:

###### 1. Krever teoretisk kompetanse

Selv om metodene i stor grad støttes av verktøy, fordrer disse metodene at man sitter med kunnskap innen statistikk for å kunne anvende og fortolke resultatene. Selv om metodene er veldokumentert og algoritmene som ligger til grunn for utformingen av modellene er tilgjengelig, stiller disse metodene krav om at man besitter estimeringsekspertise

##### Styrker relatert til et kontekstavhengig metodevalg:

###### 1. Estimeringsmodellen tilpasses organisasjonskontekst

Styrken til disse metodene er at man ut fra data man vurderer til å være representative i den gitte kontekst får utformet en estimeringsmodell som ut fra statistiske vurderingskriterier skal gi treffsikre estimater.

#### 4.4.5. Sammensatte metoder

Briand og Wieczorek har plassert analogibasert estimering her, mens jeg har valgt å skille ut dette som egen kategori. Den andre sammensatte metoden de evaluerer er hybridmetoden COBRA som eksplisitt søker å kombinere bruk av modeller og ekspertvurderinger innen en metode. Jeg tilpasser deres kategorisering ved også å inkludere metoder med anvendelse av kunstig intelligens (Finnie, Wittig et al. 1997) i denne kategorien.

Disse metodene er interessante siden de søker å koble styrkene som henholdsvis modellene og ekspertene har. Det kan likevel stilles spørsmål ved om kompleksiteten og usikkerheten i denne type metoder blir for stor. Det trengs en mer utbredt anvendelse av COBRA og lansering av alternative metoder innen denne kategorien før man kan si noe om dette. COBRA ble lansert av i 1998 (Briand, El Emam et al. 1998), men jeg har ikke funnet andre publikasjoner enn (Ruhe, Jeffery et al. 2003) som har belyst denne metoden nærmere. Hun tilpasser metoden til anvendelse innen WEB-



#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

---

prosjekter. Anvendelse av kunstig intelligens er også kommet for kort til at man har sett om de er praktisk anvendbare.

##### **Svakheter relatert til et kontekstavhengig metodevalg:**

###### **1. Høy igangskjøringskostnad**

COBRA har relatert til andre metoder en ressurskrevende. Terskelen for å ta i bruk avanserte støtteverktøy som bygger på kunstig intelligens må etter min oppfatning også karakteriseres som høy og forbeholdt virksomheter med spesielle behov eller kompetanse.

###### **2. Krav til både informasjon og ekspertise**

Skal man kombinere modeller og ekspertvurderinger er man avhengig av å ha tilgjengelighet på informasjon på det nivået modellen krever, samt at eksperter må være tilgjengelige.

##### **Styrker relatert til et kontekstavhengig metodevalg:**

###### **1. Koble styrkene i to ulike tilnærminger**

Briand mener en stor styrke med sammensatte metoder ligger i at de forsøker å ta med seg det beste fra henholdsvis ekspertvurderinger og modellbaserte metoder, og forene disse i en hybridmodell (Briand, El Emam et al. 1998).

#### **4.5. Analogibaserte metoder**

For å ha analogibasert estimering som en valgbar metode, må man sitte med et tilstrekkelig historisk grunnlag av relevante prosjekter. Denne prosjektbasen må ajourføres ikke bare ved å tilføye nye prosjekter, men også ved at man tar bort utdaterte arbeider (Shepperd, Shofield et al. 1996). Nettopp mangelen på strukturert informasjon om de historiske prosjektene er en begrensende faktor for analogi bruk (Heemstra 1992). Å bygge opp og ajourføre en prosjektdatabase fordrer ressurser.

Jeg har valgt å ikke sub-kategorisere de analogibaserte metodene da denne form for analogibruk utgjør en relativt ensartet og liten gruppe metoder. Det er utviklet ulike verktøy som kan automatisere utvelgelsen av analogiprojekter fra prosjektbasen. ESTOR (Mukhopadhyay, Vicinanza et al. 1992), ANGEL (Shepperd, Shofield et al. 1996) and ACE (Walkerden and Jeffery 1999) er eksempler på dette. Benyttes en ekspert til å velge ut analogi, må vedkommende ha detaljert kunnskap om samlingen av analogiprojekter man har data på.

En studie viser at eksperter er bedre enn verktøy til å velge ut analogier når antall analogier man har i prosjektdatabasen er lite (Walkerden and Jeffery 1999). Dette er likevel omdiskutert. Eksperter har en tendens til å velge et av de prosjekter de nylig har vært igjennom fremfor å gjøre en objektiv vurdering av et større antall analogier er påpekt (Jørgensen, Sjøberg et al. 2000).

##### **Svakheter relatert til et kontekstavhengig metodevalg:**

###### **1. Avhengig av data fra gjennomførte prosjekter**

Analogibasert estimering er helt avhengig av at man har detaljert informasjon med en tilstrekkelig kvalitet på relevante prosjekter. Forskning har vist at man bør ha minst ti relevante analogikandidater for at man skal oppnå gode og sikre estimater ved bruk av

#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

---

analogi (Shepperd and Schofield 1997). Få virksomheter har etablert rutiner for dette, og igangkjøringskostnaden med å få etablert denne blir relativt høy.

##### 2. Svak på usikkerheten til estimatet

Mens andre estimeringsmetoder kan operere med estimatintervaller med nærmere angitte usikkerhetsvurderinger, er dette vanskelig å kvantifisere når man baserer estimatene på analogier. (Walkerden and Jeffery 1997).

##### 3. Ustabilitet

Analogibaserte metoder er sårbare for feil. Selv om de tidligere refererte studier viser til meget lovende resultater, kan disse metodene anvendt på svake informasjonsgrunnlag gi virkelig svake estimater.

*"Analogy performed best in 60 % of the cases reported in published studies. In 30 % of the cases it showed the worst predictive accuracy, hence suggesting some instability. " (Briand and Wieczorek 2002) side 42*

##### **Styrkende relatert til et kontekstavhengig metodevalg:**

##### 1. Lett å bruke og gir stor tillit

Et argument for å velge analogibaserte metoder er at de er brukervennlige både med hensyn til anvendelse og i fortolkning av estimatene (Walkerden and Jeffery 1999). Metoden utmerker seg ved å være lett forståelig og tillitskapende for brukerne i utviklingsorganisasjonen ettersom den baserer seg på virksomhetens egne prosjekter (Shepperd, Shofield et al. 1996).

##### 2. Egnet ved stabile forhold

I tilfeller der man har et stabilt teknologisk miljø og tilgjengelighet på et tilstrekkelig mengde historiske data er analogibaserte metoder svært lovende (Hill, Thomas et al. 2000).

##### 3. Tidlig tidspunkt

Så snart du kan karakterisere det nye utviklingsoppdraget, kan de analogibaserte metodene gi komplette estimater på et detaljert nivå. Metodene er langt fra så spesifikke i beskrivelsene av hvilken input de krever som de modellbaserte metodene (Walkerden and Jeffery 1999).

#### **4.6. Ekspertvurderinger**

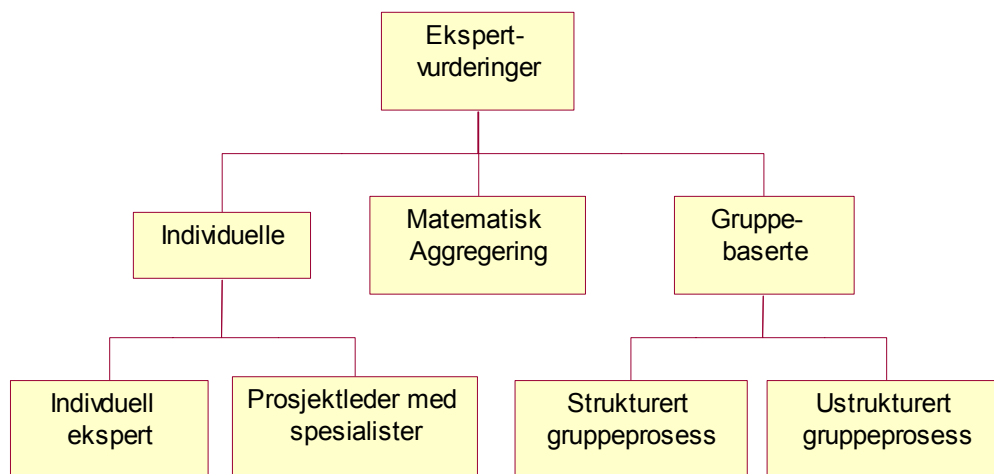
Det er når vi kommer til ekspertvurderinger at etablering av en sub-kategorisering blir virkelig utfordrende. Ekspertvurdering har som nevnt vært et nedprioritert felt innen forskning på estimering av systemutviklingsprosjekter. Når det gjelder forslag til kategorisering av de alternative metoder innen ekspertvurderinger er dette nærmest fraværende innen systemutvikling, selv om Boehm tidlig lanserte Wide-band Delphi technique som et gruppebasert alternativ til de individuelle ekspertestimatene (Boehm 1981).

Jeg har funnet en publisasjon der ekspertvurderingene kategoriseres. Her kategoriseres metodene etter om man har eksterne eller interne estimeringsgrupper, og etter om man følger en strukturerte eller ustrukturerte tilnærming (Moløkken and Jørgensen 2003). Jeg har på bakgrunn av litteratur som ser på ekspertvurderinger innen andre fagfelter og

#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

en gjennomgang av artikler som tar opp relaterte temaer innen systemutvikling, lansert et mer detaljert forlag til inndeling. Dette er gjort for å kunne bruke rammeverket på en måte som ivaretar det mangfold av tilnærminger man i praksis har innen ekspertvurderinger. Det understrekes igjen at forslaget er lansert i tilknytning til mitt rammeverk for valg av estimeringsmetode, og at det derfor ikke er modent for en mer generell anvendelse.

Jeg foreslår følgende sub-kategorisering og inndeling av ekspertvurderinger :



##### Tverrgående kategorier :

- **Top-Down / Bottom-Up strategi**
- **Bruk av sjekklister**
- **Bruk av maler**
- **Vurdering av Eksperttype**
- **Objektive kontra Subjektive eksperter**

Figur 4.2. Kategorisering av ekspertvurderinger

Hovedpoenget med forslaget er å få frem at ekspertvurderinger ikke er en ”sekkepost” av mystiske og ustrukturerte tilnærminger. Av figuren fremgår det at jeg under den hierarkiske strukturen trekker inn fem karakteristika ved ekspertvurderinger som gjør seg gjeldende på tvers av selve metodehierarkiet. Disse er i ulik grad aktuelle innen de ulike sub-kategoriene, og er viktig å ta med i vurderingen når metodevalget skal gjøres. Hva som ligger i dette beskrives nærmere etter at hver av sub-kategoriene er gjennomgått.

##### 4.6.1. Generelle svakheter og styrker ved ekspertvurderinger

Som for de modellbaserte metodene, er det enkle styrker og svakheter som gjør seg gjeldende for alle sub-kategoriene innen ekspertvurderinger. For å unngå å repetere dette for hver av kategoriene, trekkes de mest sentrale frem her. Den store fleksibiliteten til å tilpasse seg lokale behov og evnen til tidlig og på et spinkelt

#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

---

informasjonsgrunnlag kunne gi rimelig gode estimerer er en stor styrke innen ekspertvurderinger (Hill, Thomas et al. 2000).

Ekspertenes kanskje største fortrinn fremfor de alternative metodene er deres evne til å ta tak i spesielle forhold innen den gitte prosjektkonteksten og la dette innvirke på estimatene. Så vel analogier som modeller er i utstrakt grad basert på at man ut fra historikken kan si noe om fremtiden. Ekspertene er svært fleksible med hensyn til å håndtere nye tendenser og spesielle unntak som gjelder i den gitte prosjekt- og organisasjonskonteksten blant annet ved at de er gode til å utnytte sine domenekunnskaper (Jørgensen 2002).

Mangelen på dokumenterte og repeterbare prosesser har vært en akilleshæl for ekspertvurderingene i forhold til å bli akseptert som en fornuftig tilnærming til estimering (Walkerden and Jeffery 1997). Man må akseptere en stor grad av avhengighet av enkeltpersoner, og deres evne til å vurdere aktuelle situasjoner på en god måte. Ekspertvurderinger har en natur som ikke er mulig å automatisere. Metoden egner seg derfor etter min vurdering dårligere i situasjoner preget av hyppige gjentakelser og revideringer av estimeringsarbeidet.

##### **4.6.2. Individuell ekspert**

Her er det snakk om at en person som av omgivelsene vurderes til å sitte med tilstrekkelig ekspertise, blir satt til å utforme et estimat for innsatsbehovet ved et utviklingsprosjekt. Et sentralt punkt i forbindelse med valg av denne estimeringsformen er hvorvidt det er realistisk å forvente at et enkelt individ sitter med den nødvendige kompetanse til å ivareta de komplekse omgivelsene i et systemutviklingsprosjekt.

##### **Svakheter relatert til et kontekstavhengig metodevalg:**

###### **1. Skjevhetstendenser**

Ekspertvurderinger er sårbare i forhold til individuelle skjevhetstendenser og overfor politisk press og maktutfoldelse (Lederer, Mirani et al. 1990; Hughes 1996; Jørgensen 2002). Spesielt kommer denne problematikken til uttrykk innen de individuelle ekspertestimatene, da man i større grupper motvirker de individuelle store avvikene gjennom å operere med snittverdier, eller gjennom diskusjoner som minsker tendensen. Selv om ekspertvurderinger ofte treffer rimelig bra, opplever man her også de mest graverende avvikene. I et eksperiment som involverte fem eksperter, viste fire av disse seg ga rimelig gode estimerer, mens den siste ga estimerer med svært dårlig treffsikkerhet (Vicinanza, Mukhopadhyay et al. 1991). Et problem når man benytter en individuell ekspert er at man ikke vet om dette ene individet har en uheldig skjevhetstendens.

###### **2. Glemmer viktige forhold**

Systemutvikling er en kompleks oppgave. Når man legger alt ansvar på et individ og anvender en så uformell metode som ekspertvurderinger er, vil man oppleve at enkelte viktige forhold blir glemt. Dette problemet har lenge vært kjent og er stadig gjenstand for debatt (Boehm 1984; Fairley 2002). Bruk av sjekklister for å minske problemet er lansert som et virkemiddel for å minske problemet

##### 3. Krav til breddekunnskap

Det er realistisk å finne enkeltpersoner som besitter en tilstrekkelig bredde av ekspertise innen domene, teknologi og estimering når det er snakk om mindre utviklingsoppgaver. For store prosjekter mener jeg derimot at dette er urealistisk når man ser på kompleksiteten innen fagfeltet. Individuelle ekspertestimerer egner seg derfor etter min oppfatning i mindre grad i store utviklingsprosjekter.

##### **Styrker relatert til et kontekstavhengig metodevalg:**

###### 1. Lite ressurskrevende

I egenskap av å besitte ekspertise, vil vedkommende raskt og med liten ressursbruk kunne gi et estimat på innsatsbehovet. Eksperimenter der eksperter er brukt viser at ekspertene i løpet av minutter faktisk kan komme opp med estimerer av en rimelig god kvalitet, selv om dypere analyse og lenger modningstid også her må ansees som en fordel. (Vicinanza, Mukhopadhyay et al. 1991; Moløkken and Jørgensen 2003).

###### 2. Tillit

De personene som velges ut til å foreta ekspertvurderinger er som belyst i (Meyer and Booker 1991) individer som av omgivelsene betraktes som meget kompetente. Eksperter som tidligere har vist god evne til å estimere oppgaver vil ha opparbeidet seg stor tillit i egen organisasjon.

##### **4.6.3. Prosjektleder med spesialister**

Jeg har et klart inntrykk at en utbredt praksis innen systemutvikling er at prosjektleder benytter seg av spesialister som gir sine individuelle estimerer på enkeltdeler av den totale oppgaven. Prosjektlederen sitter her med et koordineringsansvar. En slik arbeidsform blir blant annet omtalt i (Jørgensen, Sjøberg et al. 2000), men jeg har ikke vært i stand til å finne publikasjoner som går nærmere i detalj på denne formen for ekspertvurderinger.

Her vil prosjektlederen be om innspill til estimerer på enkeltoppgaver av personer med spesialistkompetanse innen det aktuelle området. Som jeg vil kommentere nærmere senere, blir denne estimeringsformen preget av en Bottom-Up strategi. Prosjektleder må gjøre en vurdering av de samlede innspill som er gitt, og gi et totalestimat på prosjektet. Prosjektleder må ivareta aktiviteter som administrative støttefunksjoner og annen ressursbruk som spesialistene ikke har estimert. Av dette ser vi at man innen vårt fagfelt har et mer utvannet ekspertbegrep enn det vi finner i generell litteratur om ekspertvurderinger (Meyer and Booker 1991). En grunn til at slike spesialistvurderinger ser ut til å fungere bra innen systemutviklingsprosjekter, kan være et sterkt innslag av at spesialistkompetanse er nødvendig i forhold til å være i stand til å estimere de enkelte aktivitetene.

##### **Svakheter relatert til et kontekstavhengig metodevalg:**

###### 1. Glemmer aktiviteter

Som for de individuelle ekspertvurderingene er det her også fare for at noen oppgaver blir uteglemt når man skal estimere. Har ikke prosjektleder tilgang til en god inndeling av den totale oppgaven i form av en representativ WBS, er også denne estimeringsmetoden sårbar (Jørgensen 2003). Jeg vurderer likevel sjansene for å glemme å trekke inn vesentlige forhold under estimeringen som mindre her, da hver spesialist bidrar innen et felt der vedkommende har en god oversikt.

##### 2. Suboptimalisering

En risiko ved denne estimeringsformen er at hver enkelt spesialist gir et optimistisk estimat. Boehm har omtalt mennesker iboende tendens mot å estimere optimistisk (Boehm 1984), og hvis et antall spesialister hver for seg er optimistiske, vil det endelige estimatet bli svært skjevt.

##### **Styrker relatert til et kontekstavhengig metodevalg:**

###### 1. Reell ekspertise gjør jobben

Denne måte å estimere på er en arbeidsform som muliggjør at de som virkelig har forutsetninger for å si noe om innsatsbehovet foretar estimeringen, samtidig med at kravet om ekspertise til det enkelte individ ikke blir urealistisk. I forhold til diskusjonen om eierskap til estimatene, vil det med en slik arbeidsform ofte være personer som selv skal delta i utviklingen av programvaren som prosjektlederen velger å involvere. Et av prinsippene som Lederer foreslår er nettopp å la de som skal forestå utviklingen selv gjør estimeringen (Lederer and Prasad 1992). Ved å delegere estimeringsansvaret til de spesialistene som selv senere skal gjennomføre utviklingen, vil man kunne oppnå et økt eierskap til og en større prestisje i å holde seg til de estimatene som er satt på oppgaven.

##### **4.6.4. Matematisk aggregering**

Individuelle skjevheter er både vanskelig å unngå og vanskelig å oppdage på et tidlig tidspunkt. En måte å redusere innslaget av ekstreme feilestimeringer på oppnår man ved å la flere eksperter å foreta estimering på samme oppgave. Ferrell beskriver ulike måter å kombinere individuelle ekspertvurderinger på, og omtaler matematisk aggregering basert på henholdsvis ren gjennomsnittsberegning eller på en vektet aggregering som aktuelle metoder. Gjennomsnittsberegning er enkelt å gjennomføre og kan vise til gode resultater, mens vektet aggregering vurderes til å passe best i situasjoner der ekspertene har ulike former for ekspertise og de er vant til å jobbe sammen. Det anbefales en gruppestørrelse på tre til fem personer (Ferrell 1985). Denne formen for ekspertestimering er relevant også innen systemutvikling, og nevnes så tidlig som hos (Boehm 1981) selv om omfanget av publikasjoner som har sett nærmere på dette er minimal.

##### **Svakheter relatert til et kontekstavhengig metodevalg:**

###### 1. Ressurskrevende

Det er åpenbart at det er mer ressurskrevende å la flere eksperter forta estimering av samme oppgave parallelt. Her kommer man tilbake til det forholdet at estimering er en støtteaktivitet med et begrenset omfang av ressurser. Det er nødvendig med en nøye vurdering av om ekstrakostnaden kan forsvares ved at kvaliteten på estimatene kan forventes å bli bedre.

###### 2. Økt krav om tilgjengelighet

Ekspertene er en knapp ressurs i utgangspunktet, og i den grad man skal ha flere eksperter som gir estimater på samme oppgave, er man avhengig av at disse er tilgjengelig på det aktuelle tidspunkt. Denne svakheten gjør seg enda mer gjeldene i tilfeller der man ønsker å ajourføre estimatene utover i prosjektførløpet.

##### **Styrker relatert til et kontekstavhengig metodevalg:**

###### **1. Minsker skjevhetsproblematikken**

Et eksperiment på Personal Software Process viser at gjennomsnittet av individuelle ekspertestimerer gir god treffsikkerhet (Höst and Wohlin 1998). Setter man flere eksperter til å estimere samme oppgave, vil man redusere problemet med at enkeltindivider iblant kan gi ekstremt dårlige estimater.

##### **4.6.5. Generelt om gruppebaserte ekspertvurderinger**

Et alternativ til at enkeltindivider hver for seg sitter og estimerer større eller mindre deler av en systemutviklingsoppgave, er å la dem gjennomføre en prosess der de arbeider i gruppe. Som tidligere nevnt ble dette tidlig lansert som et alternativ (Boehm 1984), men siden ikke vært gjenstand for utstrakt forskning. For meg ser det ut til at man har vært mer opptatt av å påpeke svakheter enn av å se potensialet som ligger i denne arbeidsformen.

*“Estimation as a group activity would appear to have some advantages as the enthusiasm of some can be balanced by the caution of others, but there are dangers that people with stronger personalities may dominate the deliberations. “ (Hughes 1996), side 69*

Fremfor å være bekymret for enkeltpersoners potensielle utilbørlige dominans i en gruppe, burde man etter min mening i større grad forsket på arbeidsformen. Som diskusjonen forut har vist har alle metoder sine sterke og svake sider, og gruppebasert estimering er ikke noe unntak. Moløkken (Moløkken and Jørgensen 2003) argumenterer for at kritikken mot svakhetene i gruppeprosessene dels har vært preget av misoppfatninger, og mener denne tilnærmingen har verdier man ikke i tilstrekkelig grad har vært villig til å prøve ut.

##### **4.6.6. Strukturert gruppeprosess**

I den grad gruppeprosesser har vært omtalt, har dette vært fokusert rundt strukturerte og akademisk forankrede arbeidsformer. Ved siden av Delphi (Boehm 1984) har jeg funnet omtaler av en gruppeprosess kalt Estimeeting (Taff, Borchering et al. 1991; Walkerden and Jeffery 1997). Mens Wideband Delphi tar utgangspunkt i eksperter som gir hver sine totalestimerer for senere å komme til konsensus rundt et estimat, er Estimeeting bygd opp rundt at spesialister gir sin delestimerer, dog etter å ha vært gjennom en skikkelig gruppediskusjon rundt utviklingsoppgaven.

##### **Svakheter relatert til et kontekstavhengig metodevalg:**

###### **1. Store krav til ressurser og tilgjengelighet**

Denne arbeidsformen gir en reflektert og grundig estimeringsprosess, men er mer ressurskrevende enn alternative tilnærminger. Man må ha tilgang på et større antall eksperter, og man er avhengig av at disse er tilgjengelig på samme tid slik at møtene kan holdes. En slik arbeidsform etter min vurdering mest aktuell i store prosjektene, der enklere former for ekspertvurderinger blir vurdert som utilstrekkelig.

**Styrker relatert til et kontekstavhengig metodevalg:****1. Redusere skjevhetstendens**

Strukturerte gruppeprosessen kan i egenskap av sin formaliserte tilnærming bidra til å redusere risikoen for den uønskede skjevhetstendensen man finner i individuell tilnærminger. Hvor egnet de er til dette har som nevnt vært gjenstand for diskusjoner i det akademiske miljøet

**2. Motiverende**

Gjennom gruppeprosesser vil ekspertene treffe faglig likesinnede for interessante diskusjoner. Dette er et perspektiv som ikke bør undervurderes. Estimeringsoppgaven vil kunne bli mer motiverende, og gruppene vil kunne fungere som et utvekslingsforum for ekspertene med betydning utover det enkelte prosjektet man faktisk estimerer.

**4.6.7. Ustrukturert gruppeprosess**

Moløkken har gjennomført eksperimenter der eksperter med ulik bakgrunn blir satt til å utarbeide innsatsestimater gjennom en ustrukturert gruppediskusjon (Moløkken and Jørgensen 2003). Det er her helt sentralt av man setter sammen estimeringsgruppen med utgangspunkt i at ulike roller i utviklingsorganisasjonen er representert. I eksperimentet var henholdsvis rollene som kundekontakt, prosjektleder, bruker/analytiker/deigner og programmerer representert i estimeringsgruppen. Dette bidrar til at man gjennom en gruppediskusjon får belyst temaet fra ulike perspektiver. Ut fra de erfaringer jeg selv har, og det jeg har hørt om fra virksomheter innen bransjen, er dette en relativt utbredt estimeringsform i praksis. Dette står i sterk kontrast til at dette før Moløkkens publikasjon ser ut til å ha vært et udekket område innen forskningen.

**Svakheter relatert til et kontekstavhengig metodevalg:****1. Ressurskrevende**

Her må flere personer samles til gruppediskusjoner. Ressursene som brukes blir større enn ved individuell estimering. Slik jeg vurderer det vil dette likevel ikke ha samme omfang som de strukturerte gruppeprosessene

Ved å benytte deltagere som uansett er med i utviklingsprosjektet, vil ressursbruken kunne reduseres ved at deltagerne da i mindre grad må introduseres for problemstillingen. Dette er likevel en arbeidsform som nok først kommer til sin rett når de enklere formene for ekspertestimater ikke er tilstrekkelige, samtidig med at oppgavens art gjør modellbaserte og analogibaserte metoder lite egnet. Spesielt utfordrende prosjekter som involverer ny teknologi og nye anvendelsesområder blir trukket frem som omgivelser hvor denne type tilnærming vil komme til sin rett (Moløkken and Jørgensen 2003).

**Styrker relatert til et kontekstavhengig metodevalg:****1. Ivaretar ulike perspektiver**

Slike gruppeprosesser gir diskusjoner som er med på å sikre at man tar hensyn til totaliteten i utviklingsoppdraget på en bedre måte enn individuelle ekspertvurderinger (Moløkken and Jørgensen 2003). I forhold til de strukturerte gruppeprosessene vil denne type estimeringsprosess være mer fleksible med hensyn til å involvere de personer man har tilgjengelig på estimeringstidspunktet.



#### 4. KARTLEGGING AV ALTERNATIVE ESTIMERINGS-METODER

##### 2. Reduserer skjevhetstendens

I likhet med andre ekspertvurderinger som involverer flere eksperter, vil man også her minimere sannsynligheten for de ekstreme skjevhetene man kan oppleve ved individuelle eksperter. Mer interessant er det at denne estimeringsformen gjennom eksperimenter viser mer treffsikre estimerer enn matematisk aggregering (Moløkken and Jørgensen 2003).

##### 4.6.8. Tverrgående kriterier innen ekspertvurdering

Ved bruk av modellbaserte metoder blir de som estimerer veiledet i hvilke indikatorer de må fremskaffe. Når det gjelder de ulike formene for ekspertvurderinger er det en langt større fleksibilitet med hensyn til prosessen. Flere av de sub-kategoriene jeg her har lansert har en så stor fleksibilitet at det til dels går på bekostning av metodenes funksjon som veiledende for hvordan man skal utføre estimeringsaktiviteten. Jeg har derfor funnet det nødvendig og fornuftig å supplere den hierarkiske strukturen med et sett av karakteristika som kan bidra til å avhjelpe dette. Disse gjør seg gjeldene på tvers av de ulike sub-kategoriene.

	Individuell Ekspert	Prosjektleder med spesialister	Matematisk aggregering	Strukturert gruppeprosess	Ustrukturert gruppeprosess
Top-Down kontra Bottom-Up strategi	Både TD og BU anvendes	BU	Primært TD, Unntaksvis BU	TD (Delphi) BU(Estimeeting)	Primært BU, men også aktuelt med TD
Bruk av sjekklister	Meget aktuelt	Aktuelt	Meget aktuelt	Aktuelt	Aktuelt
Bruk av maler	Meget aktuelt	Aktuelt	Aktuelt		
Ekspertise behov (Domene, Teknologi, Estimering)	Generalist	Ulike personer dekker ulik ekspertise. Dominans av Teknologi	Generalister	Generalister	Ulike personer har ulike ekspertise
Objektive kontra subjektive eksperter	Valgbart	Primært Subjektive	Objektive	Objektive	Promært Subjektive

Tabell 4.3. Koblingene mellom kriteriene og metodene

Tabell 4.3. antyder hvordan jeg mener de tverrgående kriteriene er knyttet til metodene for ekspertvurderinger. Jeg vil presisere at denne tabellen er bare et utkast, men den viser likevel interessante perspektiver som jeg mener må trekkes inn i vurderingen av de enkelte metodenes egnethet i en gitt prosjektkontekst. Detaljbeskrivelsen som gis av kriteriene vil se nærmere på noe av dette.

##### **Top-Down kontra Bottom-Up strategi**

Ekspertvurderinger har i egenskap av sin fleksibilitet mulighet til å tilnærme seg estimeringen både gjennom Top-Down strategi og Bottom-Up strategi. Ved Top-Down strategi estimerer man utviklingsprosjektet på totalnivå. Følger man Bottom-Up strategien har, eller setter man opp en Work Breakdown Structure der den helhetlige oppgaven brytes opp i enkeltaktiviteter. Man estimerer aktivitetene enkeltvis, for så å summere opp disse til et totalestimat. Dette er en mer ressurskrevende estimeringsform. Man vil til gjengjeld sitte med både detalj- og totalestimater. I forhold til Top-Down strategien forutsetter Bottom-Up tilnærming mer detaljert informasjon om utviklingsoppgaven, noe som har innvirkning ved valg av metode.

Mange interessante aspekter ved disse to strategiene innen ekspertvurderinger belyses i (Jørgensen 2003). Jeg velger her å trekke frem at Bottom-Up strategien synes å gi en meget sterk fokus mot å estimere programmeringsaktivitetene nøye, og ut fra dette kalkulere innsatsbehovet i øvrige aktiviteter. Dette tilsier at man for å bruke denne estimeringsstrategien må besitte kompetanse innen utviklingen av løsningen. Selv om denne arbeidsformen trolig er utbredt i praksis, er det få publiserte arbeider rundt en slik tilnærming til estimering. For Top-Down strategier viser Jørgensen til at man har en annen tendens med hensyn til kravene til ekspertenes kompetanse. Her er kjennskap til relevante prosjekter for uformell bruk av analogi og kunnskap innen estimering viktigere enn teknisk innsikt innen det utviklingsarbeidet som skal gjøres.

I hovedsak presenterer litteraturen metoder for estimering av utviklingsoppgaven ut fra en Top-Down strategi (Hill, Thomas et al. 2000). I forhold til min kategorisering av ulike metoder innen ekspertvurderinger, vil det antydes i tabell 4.3. være mulig å velge mellom eller kombinere TD og BU strategiene innen flere av metodene. I de tilfeller der man har valgmuligheten, bør strategien avklares i forbindelse med at man velger metode..

Det har vært begrenset forskning på dette området og det gis ulike konklusjonene på når det er hensiktsmessig å benytte de respektive strategiene (Connolly and Dean 1997; Moløkken 2002). Jørgensen viser at tilgjengelighet på informasjon om så vel tidligere gjennomførte prosjekter som den aktuelle oppgaven har betydning. Også ekspertisen til de som skal estimere og omfanget av ressurser man vil anvende på estimeringen har bør påvirke metodevalget. Dette viser at man må gjøre en vurdering av den aktuelle konteksten før man velger mellom TD eller BU strategi. En kombinasjon av strategiene er en mulig tilnærming, men dette fordrer ytterligere ressurser til estimeringsarbeidet.

##### **Bruk av sjekklister**

Det er viktig å støtte ekspertene i arbeidet med å motvirke svakheter i deres tilnærming (Hughes 1996). Et viktig bidrag som kan virke på tvers av de lanserte kategoriene er anvendelse av sjekklister. Generelle sjekklister som dekker den totale estimeringsprosessen er lansert (Park 1996; Jørgensen 2003). I Jørgensens liste berøres også ekspertvurderinger spesielt. Det er likevel et behov for mer detaljerte sjekklister som kan støtte ekspertene. Svakheter ved ekspertene glemmer enkelt forhold, gir inkonsistente estimater og tar lite hensyn til at det er forhold de ikke kjenner til kan bli mindre ved bruk av sjekklister (Jørgensen 2002). Som tabellen viser er dette relevant innen ulike former for ekspertvurderinger.

##### Bruk av maler

For å systematisere måten ekspertene estimerer på, kan også bruk av maler være et hjelpemiddel som kan anvendes innen enkelte av tilnærmingene. Dette kan være maler som setter krav til så vel estimeringsprosessen som selve estimeringsproduktet. Poenget er som med sjekklistene å støtte ekspertene i arbeidet, og på denne måten både å bevisstgjøre ekspertene. Slike maler vil også være med på å synliggjøre estimeringsmetodene for omgivelsene, ved at man får etablert strukturer rundt ekspertene.

Et eksempel på hvordan bruk av maler kan slå ut vises i følgende sitat :

*” It turned out that this was a rather unnecessary question as the organization had a very structured and detailed template into which estimates had to fit. As will be seen, this template also appeared to be a major influence on the way that estimates were done.” (Hughes 1996), side 73.*

##### Ekspert typer

Hva man legger i begrepene ekspert og ekspertise er på ingen måte entydig innen fagfeltet. De som skal avgjøre om ekspertvurderinger er gunstig i en gitt prosjektkontekst, og i neste omgang finne ekspertene må ha et bevisst forhold til hva de ser etter. Gjennom å søke i generell ekspertvurderingslitteratur fant jeg en inndeling i ulike former for ekspertise som kan være til hjelp i denne vurderingen :

- Perseptuelle eksperter – folk som har en høy observasjonsferdighet, og som dermed ser sammenhenger andre ikke registrerer
- Kognitiv eksperter – folk med egenskaper til å forstå og løse sammenhenger andre ikke skjønner
- Eksperter med tilegnet kunnskap
- Eksperter med en utpreget vurderingsevne og en evne til å trekke slutninger
- Passive eksperter som gi informasjon til andre som må løse oppgaven
- Aktive eksperter som selv er med på å utføre beslutninger de selv har fattet

Denne inndelingen er lansert i (Shanteau 1987). Ved å kombinere denne med at man innen systemutvikling har vurdert følgende former for ekspertise som sentrale: estimeringsekspertise, domenekunnskap, utviklingsferdigheter og ledererfaring (Hughes 1996), får man et godt grunnlag for å vurdere aktuelle kandidater. I vurderingen må man ta hensyn til hvilke type ekspertise man har i organisasjonen, og eventuelt hvordan ekspertisen er spredt på ulike personer. .

I tabell 4.3. har jeg omtalt eksperter som besitter ekspertise innen teknologi, domene og estimering som *generalister*. Man kan finne individer som godt nok dekker behovet for disse ulike ekspertiseformene når det er snakk om enklere prosjekter. Innen systemutvikling har nok begrepet ekspert og anvendelsen av ekspertvurderinger blitt preget av en svært pragmatisk holdning, der man anvender de personer internt i organisasjonen som best tilfredsstiller forventningene man har til eksperter. Et annet interessant trekk er tendensen mot å la flere spesialister gå sammen om å fylle rollen som ekspert. Innen et komplisert fagfelt i rivende utvikling virker dette å være en fornuftig tilpasning av bruk av ekspertvurderinger. Dette medfører at anvendelsen av kategoriene ”Prosjektleder med spesialister” og ”Ustrukturert gruppe” har stor relevans.

Noen generell utvelgelsesmekanisme for eksperter finnes ikke. Moløkken (Moløkken and Jørgensen 2003) peker på gevinstene med å involvere eksperter med ulik bakgrunn for å betrakte prosjektet fra ulike vinkler. Hensynet til kvalitet på estimatene trekker i retning av gruppebaserte ekspertvurderinger sammensatt av eksperter som dekker de ulike typene beskrevet over. Tenker man på ressursbruk og tilgang på eksperter, vil nok likevel individuelle ekspertvurdering med støtte i etablert WBS eller sjekklister ofte være tilstrekkelig for de mindre systemutviklingsprosjektene.

##### **Objektiv kontra subjektiv eksperter**

Shanteau's inndeling subjektive kontra objektive eksperter et er forhold som har vært diskutert innen systemutviklingsforskningen uten at man har endt opp med noen klar konklusjon. Det må bli en vurdering av den aktuelle kontekst som avgjør dette. Lederer ender etter å ha diskutert ekspertvurderinger gjennom flere artikler (Lederer, Mirani et al. 1990; Lederer and Prasad 1992; Lederer and Prasad 1998) opp med at det beste rådet er å ansvarliggjøre de utviklere, ledere og estimeringsekspertene som estimerer oppgaven. Dette taler i retning av å benytte subjektive eksperter når man skal estimere.

(Londeix 1987) argumenterer derimot for å involvere en ekstern estimeringsspesialist, "Consulting Estimator". Han fraråder spesielt at prosjektlederen settes i rollen med å estimere, da han argumenterer for at vedkommende uvergelig blir overoptimistisk i estimeringsarbeidet. I (Moløkken and Jørgensen 2003) ser man på skillet som så sentralt at estimeringsmetodene kategoriseres ut fra dette kriteriet. Skal man trekke frem fordeler og ulemper her kan man si at objektive ekspertvurderinger er mindre sårbar for personlige og politiske skjevheter. Man må på den annen side forvente at organiseringen og kommunikasjon blir mer ressurskrevende her, da ekspertene i større grad må settes inn i prosjektets problemstillinger. Ved subjektive ekspertvurderinger kjenner ekspertene oppgavene i detalj, og deltagerens motivasjon kan bli positivt påvirket av at de selv blir ansvarlige for å gjennomføre den påfølgende oppgaven. Dette er forhold som den som skal velge metode må foreta en subjektiv vurdering av i forhold til de lokale forholdene man har innen et prosjekt.

## 5. KRITERIER FOR METODEVALG

Jeg har hittil argumentert for anvendelse av valgstrategien *Strukturert Vurdering* og kategorisert de alternative metodene man har å velge blant. Det siste steget for å etablere rammeverket blir å lansere de kriterier som skal benyttes ved vurderingen av de alternative metodene og koble disse til konteksten. Armstrong beskriver i tilknytning til denne valgstrategien prinsippet:

*"List the important criteria before evaluating methods". (Armstrong 2001) side369.*

I den sammenheng valgstrategien her skal anvendes må det etter min vurdering skreddersys en liste av evalueringskriterier som er særlig relevante innen systemutvikling. Jeg argumenterer for at man lager en "fagfelt"spesifikk liste. Dette betyr altså at jeg anvender Armstrongs prinsipp til å definere et generelt rammeverk til bruk i systemutvikling. Listen av kriterier skal altså ikke utformes nede på prosjektnivå, men jeg har heller ikke kunnet hente noen ferdig definert liste av valgkriterier fra andre fagfelter.

Ved utarbeidelsen av disse kriteriene er forholdet med å finne kriterier som sikrer stor innflytelse av prosjekt- og organisasjonskontekst vektlagt spesielt. Tilnærming skal ivareta at det ikke eksisterer noen "beste" estimeringsmetode, men at alternative metoder har styrker og svakheter som gjør at de passer i ulike sammenhenger (Briand and Wiczorek 2002). Anvendelse av de lanserte kriteriene i valgprosessen skal bidra til at man velger en metode som er godt egnet i de aktuelle omgivelsene. Kriteriene vil på ingen måte automatisere prosessen. I forhold til tidligere praksis vil man snarere måtte innføre en ny aktivitet. Gevinsten er at man ender opp med den estimeringsmetode man selv har vurdert til å være den riktige for oppgaven. Samtidig vil man øke bevisstheten rundt den forestående estimeringsoppgaven ved at den aktuelle prosjektkontekst innlemmes i vurderingsprosessen.

Listen av kriterier skal anvendes på tvers av det mangfold av oppgaver systemutviklere står overfor. Samtidig er det viktig at det tas hensyn til at selve valgprosessen skal gjennomføres med en minimal ressursbruk. Jeg har ut fra dette vektlagt at kriteriene må være enkle å benytte slik at et bevisst valg kan gjøres i løpet av minutter eller timer. En valgprosess som krever stor grad av informasjonsinnhenting og lang modningstid vil i praksis ikke få gjennomslagskraft i det enkelte systemutviklingsprosjekt. Dette henger sammen med knappheten på ressurser som til bruk på estimeringsaktiviteten, noe som ble nærmere belyst i kapittel 2. Hvert kriterium kobles mot de svake og sterke sidene ved metodene som ble belyst i detaljgjennomgangen i kapittel 4.

Utvelgelsen av kriteriene støtter seg på litteraturstudiene, men det endelige utvalg er også påvirket av forhold som egne erfaringer som deltager og leder av systemutviklingsprosjekter, kontakt med virksomheter som driver industriell systemutvikling og sunn fornuft. Etter først å vise hvordan jeg jobbet meg frem til et antall relevante kandidater til kriterier ved å benytte litteratur innen både prediksjon estimering av systemutviklingsprosjekter, presenterer jeg mitt skreddersydde forslag til kriterieliste til bruk i *Strukturert Vurdering*. Hvert enkelt av kriteriene blir så beskrevet

for å vise hvordan anvendelse av dette gir føringer for valg av metode som følge av den konteksten man har i prosjektet.

Jeg har forsøkt å finne støtte i publikasjoner innen systemutvikling på hvordan det enkelte kriteriet gir direkte og indirekte føringer for metodevalget. Dette har ikke vært noen enkel oppgave, siden valg av metode i så liten grad har vært beskrevet og diskutert. Dette har resultert i at noen av kriteriene er basert på et rimelig dekkende referansegrunnlag, mens andre kriterier foreløpig ikke er tilstrekkelig utforsket til at jeg kan utdype konsekvensene for metodevalg. Det blir spesielt viktig å se kriteriene i forhold til min gjennomgang av metodenes styrker og svakheter i kapittel 4.

Jeg har til tross for at referansegrunnlaget enkelte steder er mangelfullt, tilstrevet å gjøre kriterielisten mest mulig dekkende ut fra målsettingen om at rammeverket dekke helheten. Rammeverket må som det ble presisert under målsetting sees på som et forslag som skal inspirerer til nærmere utdyping og utforskning. Gjennomgangen av hvert kriterium avsluttes med at jeg gir et konkret råd til den som anvender rammeverket relatert til dette kriteriet.

### 5.1. Kriterier som tidligere er omtalt i publisert litteratur

Ettersom valg av estimeringsmetode har vært lite fokusert innen systemutvikling, og treffsikkerhet i så stor grad har eneste kvalitetskriterium for metodene innen fagfeltet, måtte jeg søke støtte i prediksjonslitteratur fra andre fagfelt for å finne inspirasjon til et mer dekkende innhold i listen med seleksjonskriterier. Jeg fant at temaet her har vært diskutert i flere tiår. Et ofte referert bidrag i denne sammenheng er artikkelen "How to choose the right forecasting technique" av Chambers m.fl. (Chambers, Mullick et al. 1971). Her påpekes viktigheten av at leder og estimeringsansvarlig sammen velger den teknikk som skal anvendes til en prediksjonsoppgave. Stadiet man befinner seg på i livssyklusen blir fremhevet som en nøkkeldeterminant ved valg av tilnærming til prediksjon. Selv om dette arbeidet var fokusert mot produkters livssyklus relatert til mer tradisjonell vareproduksjon/salg, er mange av vurderingene også relevante innen systemutvikling. Punktene under oppsummerer forhold de trekker frem som relevante i tilknytning til valg av det de omtaler som prediksjonsteknikk:

- Konteksten estimeringen foretas i
- Relevans og tilgjengelighet på historikk
- Krav til nøyaktighet
- Tidspunkt for estimat
- Verdien på estimatet for organisasjonen
- Tilgjengelig tid til å lage et estimat

Markridakis beskriver i sin bok (Makridakis, Wheelwright et al. 1983) et bredt utvalg av ulike metoder for å foreta prediksjoner. Man påpeker behovet for å ha et antall ulike kriterier å basere seg på ved valg av metode. Treffsikkerhet blir trukket frem som det viktigste seleksjonskriteriet, men samtidig bemerkes det at dette i altfor stor grad blir brukt som eneste kriterium. Kriterier som mønstre i dataene som skal predikeres, tidsserie karakteristika, tidshorisont, kostnad ved å ta i bruk alternative metoder og hvor lett man i den aktuelle organisasjonen kan anvende den aktuelle metoden, fremheves som relevante i forbindelse med valg av metode.

Et interessant poeng som trekkes frem i boken, og som jeg kommer nærmere til i mitt forslag til seleksjonskriterier, er rekkefølgen man anvender kriteriene i.

*"Conceptually, there are several ways in which criteria for selecting and comparing forecasting methods can be organized. A common approach is to prioritize criteria according to their order of importance in practice, and, as might be expected, accuracy is given top priority.*

*Unfortunately, forecasters often terminate the evaluation process after this first criterion – accuracy – has been examined."* (Makridakis, Wheelwright et al. 1983) side 761.

Ser vi på nyere arbeider innen feltet har Armstrong laget forstående tabell som oppsummerer ulike surveys innen fagfeltet "forecasting". Denne viser en meget stor enighet blant forskere, lærere, utøvere og beslutningstagere om hvilke valgkriterier som bør vektlegges ved valg av prediksjonsmetode.

*Exhibit 3. Criteria for selection a forecasting technique  
(Scale : 1= unimportant, to 7 = important)*

Criteria	Mean Importance Rating (number of respondnig)				
	Researcher (94)	Educator (55)	Practitioner (133)	Decision Maker (40)	Average (322)
Accuracy	6,39	6,09	6,10	6,20	6,20
Timeliness in providing forecasts	5,87	5,82	5,92	5,97	5,89
Cost savings resulting from improved decisions	5,89	5,66	5,62	5,97	5,75
Ease of interpretation	5,54	5,89	5,67	5,82	5,69
Flexibility	5,54	5,35	5,63	5,85	5,58
Ease in using available data	5,59	5,52	5,44	5,79	5,54
Ease of use	5,47	5,77	5,39	5,84	5,54
Ease of implementation	5,24	5,55	5,36	5,80	5,41
Incorporating judgemental input	4,98	5,12	5,19	5,15	5,11
Reliability of confidence intervals	5,09	4,70	4,81	5,05	4,90
Development cost (computer, human, resources)	4,70	5,02	4,83	5,10	4,86
Maintainance cost (data storage, modifications)	4,71	4,75	4,73	4,72	4,73
Theoretical relevance	4,81	4,20	4,43	3,72	4,40
Ability to compare alternative policies	-	-	-	-	-
Ability to examine alternative einvironments	-	-	-	-	-
Learning	-	-	-	-	-

*Tabell 5.1. Tabell med de viktigste kriteriene ved valg av prediksjonsteknikker hentet fra (Armstrong 2001) side 369.*

Igjen ser vi at treffsikkerhet gis høyeste prioritet. Listen inneholder kriterier med klar relevans ved metodevalg innen estimering av systemutviklingsprosjekter, men relatert

til problemstillingen jeg tar opp blir også denne listen for generell til å basere seg direkte på. Jeg har isteden brukt disse tre kildene som inspirasjon til å sette opp en skreddersydd liste som sterkere fokuserer kriterier som sikrer prosjekt- og organisasjonskontekst ved valg av estimeringsmetode.

Innen vårt eget fagfelt er det etter det jeg har funnet i mine litteraturstudier ingen publikasjoner som spesifikt diskuterer valg av estimeringsmetode. En artikkeltittel som virket lovende med hensyn til kontekstspesifikke valg var Navlakhas "Choosing a Software Cost Estimation Model for Your Organization: A Case Study" (Navlakha 1990), men her diskuteres kun modellbaserte metoder, og konklusjonen er at valg av metode bør basere seg på statistiske eksperimenter i den aktuelle organisasjonskonteksten. Dette samsvarer mer med Armstrongs valgstrategi *Statistical Criteria*, som jeg i kapittel 3 argumenterte for at var lite egnet innen systemutvikling.

De som går lengst i forhold til å diskutere temaet er igjen Briand og Wiecezorek i deres bidrag til boka "Encyclopedia of software engineering" (Briand and Wiecezorek 2002). Her beskrives status innen estimering av systemutviklingsprosjekter, før de også diskuterer faktorer de mener bør påvirke valget av metode. De trekker frem at følgende kriterier bør influere på metodevalget:

Praktiske hensyn :

- Akseptansenivå for generiske modeller i organisasjonen.
- Tilgjengelighet på eksperter.
- Kapabilitet til å håndtere kompleksiteten i oppgaven.

Tekniske hensyn :

- Påtenkt anvendelse av estimatene
- Tilgjengelighet på prosjektdata
- Validiteten på tilgjengelige data
- Karakteristika ved dataene

Utover disse arbeidene er det i en rekke av de publikasjonene som inngår i litteraturstudiet omtalt enkeltforhold som påvirker valg av metode. Ved utarbeidelsen av min kriterieliste har jeg forsøkt å sette disse enkeltfunnene i sammenheng med overstående kilder.

### 5.2. Forslag til valgkriterier

Jeg foreslår følgende liste av valgkriterier til bruk i en *Strukturert Vurdering* ved valg av estimeringsmetode i systemutviklingsprosjekter:

- 1) Begrensinger og rammer for estimeringsprosessen (Hva har vi mulighet til ?)
  - a. Eksterne krav
  - b. Vurdere å kutte ut eller utsette estimeringen
- 2) Formålet med og anvendelse av estimatet (Hva ønsker vi å oppnå ?)
  - a. Motiv
  - b. Treffsikkerhet
  - c. Tidspunkt



- 3) Evne og kapasitet til å estimere (Hva er vi kompetente til ?)
  - a. Avsatte estimeringsressurser
  - b. Teoretisk kompetanse
  - c. Tilgjengelighet på eksperter
  - d. Tidligere erfaringer med ulike estimeringsmetoder
- 4) Prosjektkarakteristika (Hvilke føringer gir oppgaven ?)
  - a. Størrelse på produkt og utviklingsgruppe
  - b. Systemutviklingsmodell
- 5) Tilgjengelighet på informasjon (Hva kan vi støtte oss til ?)
  - a. Data fra gjennomførte prosjekter
  - b. Spesifikasjonsnivå

Kriteriene sorterer under fem kategorier. Det er viktig å påpeke at sekvensen på kriteriene er ikke tilfeldig, men viser en foreslått rekkefølge den ansvarlige bør gå gjennom kriteriene etter i en gitt valgsituasjon. Dette skiller seg altså fra den organisering som jeg refererte fra prediksjonslitteraturen, der listene isteden var organisert etter viktighet på seleksjonskriteriene. Grunnen til å velge denne formen er at jeg ønsker å lansere en liste som sikrer ivaretagelse av den samlede kontekst. Makridakis (Makridakis, Wheelwright et al. 1983) advarte mot ensidig å fokusere på treffsikkerhet som valgkriterium, og jeg søker å unngå en skjev vektlegging ved å presentere listen også som en prosessveileder der valgkriteriene er satt opp sekvensielt. Denne organiseringen har ført til at flere av de mest kontekstsensitive kriteriene kommer mot slutten av listen.

Ser man på kriteriene i listen, har alle temaene i ulike sammenhenger vært diskutert innen forskning på estimering i systemutvikling tidligere. Det nye i mitt forslag er at disse kriteriene skal være bestemmende for valg av metode. Tradisjonelt har dette vært forhold som man har måttet ivareta på best mulig måte i etterkant av at man ut fra erfaring, overbevisning, popularitet eller statistiske sammenstillinger av treffsikkerhet har bestemt hvilken estimeringsmetode som skal benyttes i et prosjekt. Min argumentasjon går på at man har mye å hente på bruke disse kriteriene aktivt forut for at estimeringsmetoden velges, slik at man ved å gjennomføre en *Strukturert Vurdering* kan ende opp med en egnet estimeringsmetode.

Hvert valgkriterium vil i det følgende gis en detaljert gjennomgang. Her legges vekt på å vise dette kriteriets rolle i koblingen mellom estimeringsmetode og kontekst. Gjennomgangen er konsentrert rundt hvordan det enkelte kriteriet isolert påvirker valget av metode. Den som skal velge metode må etter å ha sjekket sitt prosjekt i forhold til hvert av kriteriene i den strukturerte vurderingen, foreta en helhetsvurdering før det endelige metodevalg gjøres. Ofte vil man ut fra prosjektets karakteristika ha kriterier som gir argumenter både for og i mot de enkelte metodealternativene, og avveiningen mellom metodene blir derfor ikke triviell. En slik gjennomgang kan i mange tilfeller avdekke fordeler ved å kombinere flere metoder for å tilfredsstille spesielle utfordringene i den gitte kontekst.

Enkelte av kriteriene vil i første rekke medvirke til å ekskludere metoder, mens andre kriterier styrker aktualiteten av en spesifikk metode. Ut fra mangfoldet i systemutviklingsprosjekter gis ikke spesifikke anbefalinger, men den valgansvarlige får

et hjelpemiddel til selv å gjennomføre en vurdering. Noen av kriteriene retter seg mot utviklingsorganisasjonen som helhet, mens andre går på rent prosjektspesifikke forhold. Uansett er ivaretagelse av kontekst det sentrale punktet her.

Den første kategorien skiller seg ut fra de øvrige ved at jeg her søker å avklare de overordnede føringer som er gitt for den forestående estimeringsprosessen. Valg av metode foregår ikke som en isolert aktivitet, og jeg har funnet det riktig å ta med de føringer som er gitt av omgivelsene.

### 5.3. Detaljertbeskrivelse av kriteriene

#### **Kategori: Begrensinger og rammer for estimeringsprosessen**

##### 5.3.1. Eksterne krav

###### **Beskrivelse**

Eksterne krav til estimeringen er vanlig i systemutviklingsprosjekter. Ofte er man i en leverandørrolle overfor intern eller ekstern oppdragsgiver. Det er vanlig at prosjektlederen stilles overfor krav til hvilke estimater som må fremskaffe og når disse må gis. Hvordan disse estimatene blir utarbeidet vil ofte være opp til prosjektlederen, men i enkelte utviklingsorganisasjoner og blant enkelte kunder må man forholde seg til at det er utformet politikk, standarder og rutiner også for selve estimeringsarbeidet.

###### **Konsekvenser for vurderingen av metoder**

Konkrete eksterne krav avgrenser muligheten prosjektlederen har til å velge den estimeringsmetoden som passer utviklingsgruppen og utviklingsoppgaven best. Isolert sett kan dette oppleves uheldig, men profesjonelle utviklere må forholde seg ryddig til gitte rammer. Innledningsvis i valgprosessen bør derfor prosjektleder alltid kartlegge eksistensen av eventuelle standarder og kontraktsformuleringer som gir klare føringer for valg av estimeringsmetode.

Er det detaljerte krav om anvendelse av en spesifikk estimeringsmetode skal valgprosessen avsluttes allerede her. Fokus bør isteden rettes mot en best mulig anvendelse av den forhåndsbestemte metoden. Åpner de eksterne kravene for å velge mellom angitte alternative metoder, bør den Strukturerte Vurderingen gjennomføres, men da med ensidig fokus mot å kartlegge forholdet mellom kontekst og de reelt valgbare metodene.

###### **Referanser til publikasjoner**

Temaet er lite belyst i forskningen. Jeg har funnet vurderinger som går ut på at desto mer formalisme man har i en kontrakt mellom kunde og leverandør, desto større er mulighetene for at utviklerne forbedrer sine estimeringsferdigheter (Londeix 1995). Min vurdering er at man i utstrakt grad har etablert uformelle kulturer på hvordan man utfører estimering i virksomheter, men at det i langt mindre grad er etablert styrende standarder. Mange organisasjoner ville hatt fordel av en mer eksplisitt estimeringsprosess, noe klare eksterne rammer og krav vil medvirke til (Briand and Wieczorek 2002).

### **Råd :**

Følg de metodevalg som eventuelt er angitt i standarder og rutiner, men gi alltid en vurdering av erfaringene i etterkant av prosjektet.

### **5.3.2. Vurdere å kutte ut eller å utsette estimeringen.**

#### **Beskrivelse:**

Det er viktig å gjøre en seriøs vurdering av alternativet om ikke å estimere innsatsbehovet i et systemutviklingsprosjekt. Det kan her dreie seg om å midlertidig utsette estimeringsaktiviteten til en senere fase i prosjektet, eller å kutte det ut helt. I egenskap av å være en støtteaktivitet er det, som omtalt i kapittel 2, fullt mulig å gjennomføre systemutvikling uten noen form for estimerer. Det er derfor riktig å vurdere hvorvidt det relativt sett er riktig å benytte ressurser til estimering kontra å prioritere andre aktiviteter.

#### **Konsekvenser for vurderingen av metoder:**

Hovedreglen bør ut fra den viktige styringsinformasjonen som ligger i gode estimat, være at man skal estimere innsatsbehov. Det må ikke bli slik at man unnlater å estimere fordi det er vanskelig, eller fordi man tidligere har hatt problemer med å oppnå ønsket treffsikkerhet. Derimot kan det ofte være et poeng å utsette estimeringsarbeidet så langt dette lar seg gjøre, siden treffsikkerheten på estimatene forbedres i takt med modningen av oppgaven. Konsekvensen av å kutte eller utsette estimeringen blir at man avslutter valgprosessen for eventuelt å den opp igjen på et senere tidspunkt. Anvendelse av rammeverket har likevel hatt sin verdi, da det synliggjør det brede spekteret av alternative estimeringsmetoder, og på den måten bidrar til at økt bevisstheten rundt å velge bort estimeringsaktiviteten.

#### **Referanser til publikasjoner der dette omtales spesielt :**

En studie foretatt av Heemstra (Heemstra 1992) viste at man i 35 % i av tilfellene ikke estimerte prosjektene overhodet. Dette viser at det i praksis drives profesjonell systemutvikling uten støtte i estimerer. Muligheten til å kutte ut eller utsette estimering vurderes som en god løsning i tilfeller der man vurderer at estimeringsmetodene ikke kan gi noe godt estimat (Angelis and Stamelos 2000). En annen måte å håndtere en situasjon der man vurderer det til å være umulige å estimere ut fra at tilgjengelig informasjon, er å benytte gjennomsnittsverdien av innsatsbruken i alle relevante tidligere prosjekter. Funn viser at dette kan være en meningsfull tilnærming, og dermed et alternativ til på utsette estimeringen (Jørgensen 2003).

Følgende sitat støtter min vurdering av at hovedreglen bør være at man estimerer, men at det i enkelte tilfeller er riktig å anvende ressursene til noe annet :

*"However, this data set makes it clear that even a good estimation process will not always adequately address an atypical project. For such projects, it is more important to assess the risk of doing the project than to strive to improve estimation accuracy." (Kitchenham, Pfleeger et al. 2002) side 72.*

Lederer argumenterer for å utsette estimeringen lengst mulig. På den måten unngår man tidlige urealistiske estimerer som ved siden av å gi feil styringsinformasjon, også

gir uheldige virkninger ved å bli ankerpunkter for den senere detaljerte estimeringen. (Lederer and Prasad 1992).

### **Råd :**

Utsett estimeringen så lenge som mulig, da alle estimeringsmetodene nyttiggjør seg at informasjonsgrunnlaget blir bedre.

## **Kategori: Formål og anvendelse**

### **5.3.3. Motiv**

#### **Beskrivelse:**

Beslutningsstøtte og aktivitetsplanlegging er generelle motiver for å estimere. Med en mer detaljert beskrivelse av motivet for estimeringen, kan man få en tidlig indikasjon på hvilken estimeringsmetode som utpeker seg som mest aktuell. Vidtfavnende motiver der man ønsker estimerer som både gir overordnet beslutningsstøtte og hjelp til detaljplanleggingen er lite egnet som støtte for et målrettet metodevalg.

#### **Konsekvenser for vurderingen av metoder:**

Det vesentlige her er å se etter de overordnede motivene. For å synliggjøre konsekvensene motivet i gitte kontekster kan ha på metodevalget, har jeg her valgt å skissere tre enkle eksempler.

I. Et motiv for å estimere kan være å få et bedre beslutningsgrunnlag for å avgjøre hvorvidt en utviklingsoppgave skal igangsettes. Her blir et tidlig totalestimat på innsatsbehovet viktig. Det vil være viktigere å ikke bomme stygt enn å treffe svært nøyaktig. Siden det er en engangsjobb vil ikke omfanget av ressurser man bruker for å fremskaffe estimatet være så kritisk, og behovet for å ha en repeterbar prosess er lite.

En tidlig avklaring på et spinkelt informasjonsgrunnlag taler mot modellbaserte metoder. Behovet for å unngå et grovt feilestimat kombinert med fleksibiliteten med hensyn til ressursbruk taler for anvendelse av en av metodene for gruppebasert ekspertvurdering, fremfor en mer risikofylt individuell ekspertvurdering.

II. En annen kontekst er anbudssituasjoner der det for leverandøren er avgjørende å fremskaffe et tidlig og realistisk estimat på oppgaven. Estimaten brukes til å utarbeide et bindende tilbud, og treffsikkerheten på estimatet vurderes som viktig. Oppdrag er ofte nært beslektet med andre prosjekter leverandøren har erfaring fra. Stort tidspress er vanlig og man vil oppleve stadige gjentakelser av tilsvarende estimeringssituasjoner da andre kunder ber om tilbud på lignende løsninger.

Ut fra motivet taler tidspunktet igjen mot modellbaserte metoder. Ønske om treffsikre estimerer på et tidlig tidspunkt kombinert med det store antallet av beslektede prosjekter, gjør etablering av en prosjektdatabase med anvendelse av analogibasert estimering til en metode man bør vurdere nærmere i den videre gjennomgangen.

III. Et tredje motiv kan være at prosjektlederen ønsker estimerer for å foreta en effektiv disponering av utviklingsressursene i prosjektet. Her er det behov for detaljerte estimerer nede på aktivitetsnivå og informasjonsgrunnlaget på estimeringstidspunktet

vil være mer omfangsrikt. Her kan alle hovedformene av estimeringsmetoder være aktuelle. Skal ekspertvurdering anvendes må den være basert på en Bottom-Up strategi. Ellers er motivet for generelt til at man kan trekke ytterligere slutninger rundt metodevalg.

Av eksemplene ser man at motivene varierer med konteksten. Siden enkelte av metodene har sin styrke i å gi tidlige totalestimer av en rimelig kvalitet, mens andre metoder krever omfattende informasjon, men til gjengjeld kan gi mer treffsikre og detaljerte estimer, blir motivet styrende for metodevalg.

### **Referanser til publikasjoner der dette omtales spesielt :**

Estimering er en aktivitet med stor grad av iboende problemer og metodevalget bør sees i sammenheng med at man bør ha realistiske motiver fremfor idealistiske forventninger (Skagestein, 2002). Her bør den som skal velge metode formulere et mest mulig presist motiv for estimeringen, for på den måten å innsnevre antall aktuelle estimeringsmetoder.

Dette kriteriet viser ellers godt hvordan de styrker og svakheter jeg beskrev ved metodene i kapittel 4 skal anvendes ved gjennomgangen av de enkelte valgkriteriene. I kapittel 4 ble det lagt vekt på å finne støtte i relevant publisert forskning, så en direkte anvendelse av det jeg har beskrevet som metodenes styrker og svakheter skulle være faglig forsvarlig. Dette gjør at man får koblet samme metodekarakteristika med konteksten, noe som jo er det sentrale poenget i mitt rammeverk.

### **Råd :**

Formuler et mest mulig presist motiv, og se hvilken tendens dette gir i forhold til metodevalg.

### **5.3.4. Treffsikkerhet**

#### **Beskrivelse:**

Det dominerende kvalitetsmålet på estimeringsmetodene har innen forskningen vært treffsikkerheten på estimatene. Implisitt i dette ligger at treffsikkerheten bør være bestemmende for valg av metode. Innen systemutvikling har MMRE og Pred() etablert seg som to anerkjente måter å måle treffsikkerhet på. Disse ble foreslått av Conte m.fl.(Conte, Dunsmore et al. 1986) og er senere omtalt som utbredte kvalitetsmål for estimer av blant andre Shepperd og Kitchenham (Shepperd and Schofield 1997; Kitchenham, Pfleeger et al. 2002).

Også innen systemutvikling har man dermed gått i den fella som Markidrakakis advarte mot, nemlig å fokusere kun på det som er allment akseptert som det viktigste seleksjonskriteriet (Makridakis, Wheelwright et al. 1983). Til tross for konsentrert forskningen rundt dette ene kriteriet, har man likevel ikke klart å komme frem til enighet om hvilke metode som gir den beste kvalitet. Treffsikkerhet på estimatene hører åpenbart med som et seleksjonskriterium også i min liste, men skal ikke alene gi svar på hvilken metode man bør velge. Mekanismene for å håndtere treffsikkerhet er ulike i de alternative metoder, og kravene til treffsikkerhet i en gitt kontekst bør derfor innvirke på metodevalget.

### **Konsekvenser for vurderingen av metoder:**

Mange sammenligningsstudier er gjort på dette området, og som tidligere referert viser Jørgensens sammenstilling (Jørgensen 2002) av 15 studier at man ikke kan konkludere med at noen metode generelt gir bedre treffsikkerhet enn de andre.

Det er akseptert innen fagfeltet at intervallestimer er bedre enn punktstimer når det gjelder å gi brukerne muligheten til å vurdere treffsikkerhet. Analogibaserte metoder håndterer ikke intervaller på noen god måte. Man kunne tro at formelle modeller utledet gjennom statistiske metoder var gode på angivelse av konfidensintervaller. Publikasjoner viser derimot til at modellene ofte er svakere til dette enn ekspertene, til tross for at forskningen viser at eksperter har en klar tendens til å gjøre intervallene altfor snevre (Jørgensen 1995; Jørgensen and Moløkken 2002; Jørgensen and Teigen 2002). Siden konklusjonene i ulike studier er så vidt forskjellige, og er det vanskelig å peke øvrige spesielle tendenser her.

### **Referanser til publikasjoner der dette omtales spesielt :**

Forskningen har ikke gitt entydige svar med hensyn til hvilken metode som gir størst treffsikkerhet. Man har likevel en rekke funn som på ulik måte gir oss støtte når vi skal vurdere treffsikkerhet relatert til metodevalg.

Kalibrerte modellbaserte metoder gir høyere treffsikkerhet enn de ukalibrerte. (Kemerer 1987)

Individuelle eksperter treffer ofte bedre enn modellbaserte metoder, men er samtidig meget utsatt for de mest ekstreme feilestimatene. (Vicinanza, Mukhopadhyay et al. 1991).

Funn tyder på at man ved bruk av modellbaserte metoder er mindre utsatt for virkelig store avvik i treffsikkerheten, da modellene med sin utstrakte bruk av historiske data fjerner noe av faren for skjevhet (Jørgensen 2002).

Funn har vist at analogier har størst treffsikkerhet i 60 % av tilfellene, men samtidig er verst i 30 % (Briand and Wieczorek 2002).

Publikasjoner viser at fokus bør rettes mot å forbedre ekspertenes evne til å angi mer realistiske konfidensnivåer og å estimere den nødvendige bredde på det tilhørende minimums- og maksimums-innsatsbehovet. (Jørgensen, Teigen et al. 2002; Jørgensen 2003)

### **Råd :**

La ikke treffsikkerheten være det dominerende valgkriteriet.

### **5.3.5. Tidspunkt og hyppighet**

#### **Beskrivelse:**

Krav rundt tidspunktet for fremstillingen av estimatet er et kjerneområde i forbindelse med valg av metode for estimering. De ulike metodene er svært forskjellige med

hensyn til kravene til informasjonsgrunnlag, og tidspunktet for når estimatene må foreligge har dermed direkte innflytelse på hvilke metoder man har å velge blant.

Ved siden av krav om når det første estimatet må foreligge, påvirker også krav om ajourføring av estimatene gjennom prosjektets livssyklus metodevalget. For å fungere som styringsinformasjon gjennom utviklingsforløpet, må estimatene gjennomgås og ajourføres flere ganger i løpet av en systemutviklingsprosess. Estimeringsmetodenes ulike grad av automatisering må vurderes ut fra behovet for ajourhold.

### **Konsekvenser for vurderingen av metoder:**

Krav om svært tidlige estimater er et kriterium som begrenser antall alternative metoder for estimering. Ulike former for Top-Down baserte ekspertvurderinger er metoder som kan benyttes tidlig i et prosjekts livssyklus, da med et høyt innslag av risiko hvis informasjonsgrunnlaget er lite. Her vil en utstrakt bruk av sjekklister medvirke til at risikoen reduseres.

Så snart det er mulig å karakterisere et prosjekt vil Analogibaserte metoder være et interessant alternativ. Man vil ofte være i stand til å fremskaffe informasjon som i tilstrekkelig grad karakteriserer prosjektet på et tidlig tidspunkt, særlig hvis prosjektet er av en type man har tradisjon for å gjennomføre i virksomheten.

Majoriteten av de modellbaserte metodene vil derimot være utilgjengelige inntil flere detaljer rundt kravene til den nye løsningen er formulert. Dette trengs for at størrelse og produktivitet skal kunne predikeres med tilstrekkelig sikkerhet. Også ekspertvurderinger med en Bottom-Up strategi fordrer et informasjonsgrunnlag som først foreligger noe ut i prosjektforløpet. På denne måten ser vi at krav til tidspunktet man krever estimatet påvirker valget.

Kravene til frekvens på ajourhold påvirker metodevalget på andre måter. Innen ekspertvurderinger må man i stor grad gjenta samme prosessen som tidligere ved reestimering fordi metoden har en minimal automatiseringsgrad. Er frekvenskravet høyt går ekspertvurderingen fra å være en rimelig og fleksibel tilnærming til å bli ressurskrevende og problematisk. Spesielt gjelder dette de ulike formene for gruppebaserte ekspertvurderinger.

Analogibaserte metoder krever kun en justering av prosjektkarakteristika og eventuelt en justering av tilpassningsfaktorer. Modellbasert estimering er også velegnet ved hyppige ajourføringer ut fra den høye grad av automatisering som ligger i de fleste av disse metodene.

### **Referanser til publikasjoner der dette omtales spesielt :**

Verdien av en estimeringsmetode blir vurdert som høyere desto lavere krav den har til detaljert informasjon. Det er i de tidlige stadiene i et utviklingsprosjekt at estimatene er viktigst (Gray, MacDonnell et al. 1999). Modell-baserte metoder krever i utgangspunktet informasjon som først er tilgjengelig et stykke ut i utviklingsforløpet. Et karakteristisk eksempel på at metodene er forsøkt tilpasset problematikken er lanseringen av COCOMO II (Boehm, Clark et al. 1995). Her har man innført en tilpasningsdyktig kombinasjon av tre modeller som gjør metoden bedre egnet for bruk tidligere i et prosjekts livssyklus.

Når det gjelder kapasiteten til de analogibaserte metodene tidlig i et prosjekts faser beskrives denne som særdeles lovende og overlegen alternative metoder av (Shepperd and Schofield 1997).

Dynamikken i et systemutviklingsprosjekt krever ajourførte estimater, noe Modellbaserte metoder takler godt. Et eksempel på at man forsøker å perfektionere de modellbaserte metoder til de dynamiske omgivelsene gis i (Schooff and Haimes 1999), der man utvider og kompliserer eksisterende modeller for å ivareta dette. Jeg ser klare ulemper med å tilføye ytterligere kompleksitet i metoder for å takle dynamikken i omgivelsene. Isteden finnes støtte i bl.a. (Chatzoglou and Macaulay 1996) for at man heller bør vurdere å bytte metode etter hvert som prosjektet går gjennom sin livssyklus og kravene til estimatene endres.

### **Råd :**

Avklar tidspunkter og ajourholdsfrekvensen samtidig da disse kriteriene ofte stiller ulike krav til metodene

### **Kategori: Evne og kapasitet til å estimere**

De ulike estimeringsmetodene stiller forskjellige krav til utøverne når det gjelder kunnskaper om estimering generelt og til behovet for innsikt i detaljene om metodene spesielt. Omfanget av tid og ressurser man må regne med å bruke på estimeringen varierer mellom metodene. Disse forholdene er det viktig at det tas hensyn til ved valg av metode. Ingen er tjent med å anvende en metode man ikke har forutsetninger for å anvende på en god måte.

### **5.3.6. Avsatte ressurser til estimering**

#### **Beskrivelse:**

Ressursene man avsetter til estimering vil variere fra prosjekt til prosjekt. Prosjektledere er fokusert på å få best mulige estimater ved bruk av minst mulig ressurser, noe som i utgangspunktet vil favorisere metoder som enten er lite ressurskrevende eller som er fleksible med hensyn til anvendelse av ressurser.

Estimeringsspesialister er av flere trukket frem som interessant uavhengig av metodekategori. Muligheten til å ha en slik ressurs tilgjengelig i en organisasjon henger nøye sammen med de ressursene man ønsker å bruke på støtteaktiviteten estimering. Trolig er det bare de store aktørene med estimatene med innslag av forretningskritisk karakter som er villige til å prioritere dette

#### **Konsekvenser for vurderingen av metoder:**

De mer ressurskrevende metodene må overbevise de ansvarlige om at ekstra bruk av ressurser gir gevinster i form av bedre kvalitet på estimatene. Individuelle ekspertvurderinger kan ofte utføres i løpet av minutter, særlig i tilfeller der man har en prosjektleder som besitter den nødvendige ekspertise. Prosjektledere som benytter seg av spesialister som delansvarlige for estimeringen vil kreve større ressurser. De ulike formene for gruppebaserte ekspertvurderinger vil generelt kreve langt større ressurser til



estimeringsoppgaven. Anbefalte gruppestørrelser på tre til fem personer og gjentatt involvering av disse vil være ressurskrevende.

Når det gjelder analogibruk, må man etablere og vedlikeholde en prosjektbase som inneholder et nødvendig antall representative prosjekter. Dette krever et større initielt ressursforbruk. Det er her viktig at man vurderer ressursbruken i forhold til at prosjektbasen kan anvendes av mange prosjekter. Selve estimeringsarbeidet vil kreve minimalt med ressurser ved analogibruk. Ved modellbasert estimering vil en kalibrering av modellen til det lokale miljøet være ressurskrevende. Studier viser at bruk av modellbaserte metoder uten en slik kalibrering gir meget svak treffsikkerhet (Kemerer 1987). Bruk av estimeringsspesialister og kombinasjon av flere estimeringsmetoder vil være et positivt bidrag til å forbedre estimeringsarbeidet, men det vil alltid være snakk om et høyere ressursforbruk ved slike tilnærminger, noe det må argumenteres for.

Et interessant trekk i tilknytning til ressursbruk er metodenes ulike evne til å benytte seg av informasjon og ressurser som uansett må fremskaffes i tilknytning til utviklingsaktivitetene. Subjektiv ekspertestimering gjennomføres av personer som uansett tilhører prosjektgruppa og kjenner prosjektet, og selve estimeringsinnsatsen blir dermed minimal. Modellbaserte metoder som tar i bruk Use Case som informasjonsgrunnlag er et annet eksempel på å nyttegjøre seg informasjon som ikke primært er fremskaffet for estimering.

### **Referanser til publikasjoner der dette omtales spesielt :**

Temaet om ressursbruk er grundig diskutert i forbindelse med gjennomgangen av enkeltmetodenes sterke og svake sider i kapittel 4, og repeteres derfor ikke her. Spesialistrollen er derimot lite berørt tidligere. Denne omtales som et særdeles viktig virkemiddel av (McCulla 1989). En person i en slik rolle vil være en sentral ressurs i alle prosjektene en organisasjon gjennomfører, og dermed vil bidra sterkt til både å heve estimeringskvaliteten, men også til å innføre en bedre struktur i estimeringsprosessen (Londeix 1995).

### **Råd :**

Velg aldri en estimeringsmetode som fordrer større estimeringsressurser enn det du er innstilt på å bruke.

### **5.3.7. Teoretisk kompetanse om estimering**

#### **Beskrivelse:**

Prosjektledere med en datafaglig utdannelsesbakgrunn fra universiteter og høyskoler har ofte fått en introduksjon til temaet i generelle systemutviklingskurs. Det er ikke vanlig med kurs som går i dybden innen temaet, da estimering betraktes som en støtteaktivitet. Den skjeve fremstilling som gis om emnet i generelle lærebøker om systemutvikling vist i tabell 2.1., gir i beste fall en snever plattform i forhold til å foreta bevisste metodevalg.

#### **Konsekvenser for vurderingen av metoder:**

Kunnskap om de ulike metodene vil gi et bedre grunnlag for å vurdere dem. Teorikunnskapen fra utdanningen vil, supplert med egne erfaringer innen estimering og den kulturen som eksisterer rundt estimering i virksomheten, være det grunnlag den ansvarlige for metodevalg har å støtte seg på. Det er et mål at mitt rammeverk skal bøte på noe av mangelen på innsikt i de alternative metodene. En kjennskap til hovedtrekkene i innholdet i så vel kapittel 4 som 5 i dette hovedfaget, vil i gi en bedre plattform for å forstå valget.

Ekspertvurderinger er intuitivt lettere å forstå og sette seg inn. De har også oppnådd stor utbredelse, og der derfor kjent i de fleste virksomheter. De modellbaserte og analogibaserte metodene stiller høyere krav til teoretisk innsikt. Forskere og utviklere av verktøy for estimering har riktignok lagt vekt på å gjøre metodene mest mulig uavhengige av forkunnskaper, men en viss forståelse for prosessen som fører frem til estimatene gir brukerne større trygghet overfor metoden.

### **Referanser til publikasjoner der dette omtales spesielt :**

Jeg har ikke funnet publikasjoner som nærmere analyserer nivået på den teoretiske kunnskapen hos de som utfører estimering av systemutviklingsprosjekter. Enkelte studier antyder likevel at dette nivået generelt er lavt. Når analogibasert estimering oppgis å være estimeringsmetode, samtidig med at man ikke har data på de historisk gjennomførte prosjektene tyder dette på at man ikke har kunnskaper om hva som ligger i denne metoden (Heemstra 1992). Enkelte studier antyder at det i liten grad gis opplæring og trening i å estimere (Jørgensen and Sjøberg 2001).

Behovet for å forstå hva som ligger til grunn for de estimatene som man kommer frem til har i større grad vært omtalt.

*"It is often the case that practitioners want to understand how an estimate was obtained before relying on it. Those human factors play a very important role in the adoption of a resource estimation method." (Briand and Wiecek 2002) s. 27.*

At ekspertvurderinger gjennom å være intuitivt lett forståelige kan føre til at man undervurderer behovet for å sette seg nærmere inn i hva som ligger i denne metoden. Dette forholdet trekkes frem av (Jørgensen and Sjøberg 2001) som mener at mer teoretisk kunnskap om ekspertvurderinger samt mer utstrakt feedback ved anvendelse av metoden, vil føre til større treffsikkerhet.

### **Råd**

For de estimeringsansvarlige er opplæring innen grunnprinsippene i de alternative metodene viktigere enn kursing i detaljer innen enkeltmetoder.

### **5.3.8. Tilgjengelighet på eksperter**

#### **Beskrivelse:**

Bruk av rammeverket skal også føre til økt bevisstheten rundt hva man skal se etter for å vurdere hvem man kan betrakte som ekspert. På grunnlag av dette må man vurdere om det i det gitte prosjekt finnes ekspertise som muliggjør anvendelse av de ulike metodene. Man må altså bort fra det jeg tolker som å være en vanlig oppfatning innen fagfeltet, nemlig at prosjektlederen i egenskap av sin rolle, automatisk besitter

ekspertise og selv kan foreta estimeringsarbeidet enten alene eller gjennom å involvere noen få utvalgte spesialister i en uformell gjennomgang.

Det er ikke nok å avklare at virksomheten har eller kan leie inn personell som kan løse oppgaven. Man må også avklare om disse ofte knappe ressursene er tilgjengelig til rett tid og i nødvendig omfang. Behovet for ajourhold av estimatene vil innvirke på vurderingen av tilgjengelighet.

### **Konsekvenser for vurderingen av metoder:**

De modellbaserte og analogibaserte estimeringsmetodene vil ikke kreve direkte tilgjengelige eksperter på estimeringstidspunktet. Eksperter involveres her i forbindelse med kalibrering eller ved opprettelse av prosjektdatabaser, og selve estimeringsoppgaven kan i stor grad utføres av prosjektleder. Man blir således ikke sårbar med hensyn til tid og sted, og fravær av eksperter vil ikke hindre estimeringsarbeidet.

Alle former for ekspertvurderingen er derimot helt avhengig av at de eksperter og spesialister som skal estimere er tilgjengelig på det aktuelle tidspunkt. Knapphet på eksperter både med hensyn til eksistens og tilgjengelighet gjør at individuelle ekspertestimerer er mer realistiske å få gjennomført enn gruppebasert estimering. Det er også grunn til å hevde at en subjektiv estimeringsform lettere kan gjennomføres, da ekspertene vil være tilgjengelige i form av sitt engasjement i utviklingsgruppen.

For små oversiktlige oppgaver finnes enkeltpersoner som kan ivareta så vel teknologiske, estimeringsmessige og domenemessige forhold. Dermed er individuell ekspertvurdering en meget aktuell estimeringsform. Med økende størrelse og kompleksitet vil tilgjengeligheten på enkeltindivider som sitter med tilstrekkelig kompetanse minke, noe som aktualiserer valg av Spesialistrollen, samtidig med at det blir fornuftig å anvende kombinasjoner av Top-Down og Bottom-Up strategier.

Prosjektleder som involverer individuelle spesialister til å estimere deler av oppgaven er interessant innen systemutvikling, da man kan støtte seg til spesialister som reelt dekker kravene til ekspertkompetanse innen enkeltområder, men som ikke kan ivareta helheten. Har man adgang til å benytte flere eksperter, bør man legge vekt på å sette sammen en gruppe av komplementære individer, fremfor å ha en homogen samling eksperter med sammenfallende perspektiv.

Vektleggingen av ulike kompetanse varierer mellom de ulike estimeringsmetodene. Mens de modellbaserte metodene er bygd opp rundt estimeringsekspertise og teknologisk kompetanse, er ekspertvurderingene fleksible med hensyn til å trekke inn alle former for kunnskap som ekspertene besitter.

### **Referanser til publikasjoner der dette omtales spesielt :**

Et interessant perspektiv for å vurdere mulighetene til å foreta gode ekspertvurderinger er å se nærmere på hvilken type ekspertise man har tilgjengelig. Referanser innen emnet finnes i kapittel 4 der jeg viste til Shanteu og Hughes. Først etter å ha kartlagt eksistensen av og tilgjengeligheten på ekspertise i den gitte kontekst, er man i stand til å gjøre en vurdering av om organisasjonen overhode sitter med ressurser til å gjennomføre ekspertvurderinger på en god måte.

Moløkken argumenterer for at ustrukturert gruppeestimering gir et innslag av avklarende og berikende diskusjoner hvis man oppretter grupper sammensatt av personer som har ulike roller i en virksomhet (Moløkken and Jørgensen 2003).

Man har funnet en klar tendens i retning av at ekspertvurderinger gir en større treffsikkerhet på estimatene enn de modellbaserte metodene i situasjoner der ekspertene besitter og anvender domenekunnskap ved estimeringen. (Jørgensen 2002).

### **Råd :**

Velg estimeringsmetode først etter å ha vurdert den reelle tilgjengeligheten på ekspertise i den gitte prosjektkontekst.

### **5.3.9. Tidligere erfaringer med ulike estimeringsmetoder**

#### **Beskrivelse:**

Erfaringene man gjør ved anvendelse av de ulike estimeringsmetoder innen en virksomhet bør systematiseres og anvendes i fremtidige valgsituasjoner. Ved siden av spesielle karaktertrekk for prosjektet har jeg understreket betydningene av å ivareta den organisasjonsmessige konteksten. Metoder som har vært benyttet innen virksomheten med suksess i tidligere prosjekter bør benyttes i senere prosjekter, da man her har opparbeidet tillit basert på erfaringer. Tidligere suksess er dog ikke ensbetydende med at dette gjentar seg. Tilsvarende bør metoder som absolutt ikke fungerte i organisasjonskonteksten i utgangspunktet unngås da en mulig årsak til at man feilet er at metoden ikke passer den gitte kontekst.

#### **Konsekvenser for vurderingen av metoder:**

Dette kriteriet tilsier at man skal bruke det man har erfart at virker i organisasjonen tidligere. Når det gjelder de ulike metodene, vil trolig de organisasjonsmessige føringene i retning av å holde seg til en etablert praksis være sterkest der det er brukt ressurser på å bygge opp en prosjektbase for anvendelse av analogibaserte estimering, eller der man har kalibrert en modellbasert metode for bruk i organisasjonen.

Gevinster i form av repeterbarhet er langt mindre innen ekspertvurderinger, og erfaringene med å bruke metoden er langt mer personavhengig. Her er det derimot svært viktig å kartlegge enkeltpersoners ferdigheter når det gjelder å foreta estimering. Å identifisere de individer som er gode til å gi estimerer er en stor utfordring, og det kanskje beste vurderingsgrunnlaget her er resultatene fra deres tidligere estimeringer.

#### **Referanser til publikasjoner der dette omtales spesielt :**

Et godt råd er at man bør velge eksperter som kan vise til gode resultater når det gjelder å gi estimerer. Å estimere er svært vanskelig, og det er langt fra noen selvfølge at den som vurderes til å være den største ekspertern gir de beste estimatene (Jørgensen 2002). Jørgensen anbefaler derfor at man i en oversikt over kompetansen til ekspertene skiller mellom det han kaller "know-how; know-how-much; know-how-uncertain". I gruppebasert estimering har man mulighet til å koble inn eksperter med ulike ferdigheter (Moløkken and Jørgensen 2003).

**Råd:**

Bruk metoder du tidligere har hatt suksess med, og eksperter som tidligere har gitt gode estimater.

**Kategori: Prosjekt karakteristika**

Før jeg ser nærmere på de prosjekt karakteristika jeg mener har størst betydning for metodevalget, kommenteres hvorfor "Applikasjons- og oppgavetype" er utelatt. Rent intuitivt var dette et kriterium jeg mente hørte hjemme her. En nærmere gjennomgang av publikasjoner med hensyn til hvilke føringer dette gir for metodevalget, ga ikke nok holdepunkter til å fremme dette som et eget valgkriterium.

**5.3.10. Størrelse på produkt og utviklingsgruppe****Beskrivelse:**

Omfanget og kompleksiteten på systemutviklingsprosjekter varierer sterkt. Man har alle typer størrelser fra prosjekter med varighet på noen få uker og med få involverte, til prosjekter som går over flere år, der tusenvis av månedsverk brukes på utviklingen. Estimeringsmetodene er ikke spesialisert til å takle utviklingsoppgaver av en nærmere angitt størrelse, men jeg vil argumentere for at metodenes ulike styrker og svakheter også har innflytelse på hvor godt de håndterer størrelsesproblematikken.

**Konsekvenser for vurderingen av metoder:**

Før jeg ser nærmere på de enkelte metodene er det naturlig å nevne at det for de store prosjektene er meget aktuelt å kombinere flere estimeringsmetoder. Konsekvensene av å operere med feilestimer her kan bli meget stor, og villigheten til å benytte større ressurser på selve estimeringsprosessen er tilstede siden prosjektleder her har et stort behov for god styringsinformasjon.

Det er funnet støtte for at ekspertvurderinger er godt egnet ved mindre prosjekter, mens ekspertenes evne til å takle komplekse og store prosjekter er relativt svak. Spesielt gjelder dette de individuelle estimeringsformene, der det er vanskelig for enkeltindivider å overskue, ha innsikt i og evne til å ivareta den totale kompleksiteten i oppgaven. Bruk av sjekklister og maler kan avhjelpe noe, og en Bottom-Up strategi kan være med på å bedre ivaretagelse av helheten. Ensidig bruk av ekspertestimering er likevel uheldig i slike prosjekter, da risikoen for grov feilestimering er så vidt stor og konsekvensene av feilestimering tilsvarende.

Når det gjelder analogibruk er det tilgjengeligheten på relevante analogier som er avgjørende. For de helt store prosjektene minsker mulighetene til å finne analogier dramatisk, da dette er en type prosjekter som sjelden gjennomføres. Modellbaserte metoder er bygd for å mestre også de store prosjektene. I store prosjekter vil også forholdene være tilrettelagt for å anvende dem. Få virksomheter setter i gang store prosjekter med mange involverte uten først å spesifisere oppgaven relativt detaljert. Et annet forhold som taler til de modellbaserte metodenes fordel er at man i store prosjekter i større grad også vil sitte med personer med spesifikk estimeringskompetanse, noe som vil minske problematikken med at disse metodene er vanskeligere å bruke enn ekspertvurderinger.

Jeg argumentere ut fra dette for at man bør tendere mot ekspertvurderinger i små prosjekter der ekspertene lett kan skaffe seg oversikt. Analogibaserte metoder er mest egnet for det som er den typiske prosjektstørrelsen innen en virksomhet, mens modellbaserte metoder og kombinasjon av flere estimeringsmetoder kommer til sin rett i store utviklingsprosjekter.

### **Referanser til publikasjoner der dette omtales spesielt :**

Også knytningen mellom størrelse og metodenes styrker og svakheter ble i stor grad belyst i kapittel 4. Det finnes støtte for å si at ekspertene ikke er særlig gode til å estimere store oppgaver, og at en viktig årsak til at de har en tendens til å underestimere slike oppgaver ligger i at de ikke forstår den samlede kompleksiteten i store utviklingsoppgaver (Hill, Thomas et al. 2000). Tendensen mot at eksperter systematisk underestimerer store og komplekse oppgaver har sitt motstykke i de samtidig ofte overestimerer dekomponerte, små og enkle oppgaver (Connolly and Dean 1997; Gray, MacDonnell et al. 1999). Når det gjelder de små oppgavene viser så vel modellbaserte metoder som Top-Down basert individuell ekspertvurdering rimelig god treffsikkerhet. Har man tilgang på en individuell ekspert, er det naturlig å følge prinsippet om å bruke den enklest mulige metoden som gir et godt nok resultat (Armstrong 2001).

Innen de modellbaserte metoder viser for eksempel de ulike modusene i COCOMO at man har lagt stor vekt på at metoden skal takle applikasjoner av ulik størrelse. Her søker man å løse dette ved å lage varianter av modellen der Organic er tilstrekkelig i små prosjekter, Semi-detached egner seg i noe større utviklingsoppdrag mens Embedded skal takle kompleksiteten i de store utviklingsprosjektene. (Boehm 1984). Jeg har ikke funnet publikasjoner som mer detaljert sier noe om hvilken av de modellbaserte metodene som best takler store komplekse utviklingsoppgaver.

### **Råd :**

Velg den enkleste estimeringsmetoden som kan mestre kompleksiteten og størrelsen på oppgaven.

### **5.3.11. Systemutviklingsmodell**

#### **Beskrivelse:**

Estimering foregår ikke som noen isolert aktivitet. For å estimere en utviklingsoppgave er man avhengig av innblikk i det som skal gjøres, og estimering blir på denne måten en integrert del av systemutviklingen. Man er følgelig nødt til å forholde seg til den arbeidsform man har valgt å legge opp det samlede utviklingsprosjektet etter. Her kommer systemutviklingsmodellen prosjektet følger inn i bildet. Fra å ha tradisjon for å følge en sekvensiell prosess i form av en fossefallsmodell, er fagfeltet nå preget av at det finnes en rekke tilnærminger, der innslaget av iterative og inkrementelle prosesser er blitt stadig mer fremtredene.

Fossefallsmodellen ble etter hvert supplert med evolusjonære utviklingsmodeller, spiralsmodellen og senere inkrementelle og iterative utviklingsmodeller. Unified Process er en utbredt systemutviklingsmodell som søker å forene iterasjoner og inkremitter. eXtreme Programming (XP) er lansert som en tilnærming til

utviklingsoppgaven som søker å ivareta dynamikken i utviklingsarbeidet ved anvendelse av et sett praksiser som setter ramme for arbeidsformen.

### **Konsekvenser for vurderingen av metoder:**

Svært tidlig estimering påvirkes i liten grad av utviklingsmodellen da estimering her gjøres før utviklingsmodellen i særlig grad preger arbeidet. Estimering lengre ut i prosessen blir i langt større grad vil påvirket av utviklingsmodellen da spesielt av modellenes fokus på kartlegging av kravene til systemet. De modellbaserte estimeringsmodellene ble i stor grad lansert i perioden da fossefallsmodellen var dominerende, og egner fortsatt best i omgivelser preget av tidlig dekkende spesifisering av oppgaven.

Moderne utviklingsmodeller gir i økende grad føringer for estimeringsarbeidet. Unified Process har en disiplin for prosjektstyring som også innebærer råd om estimering (Kruchten 2000). Det understrekes det at den eneste måte å utforme troverdige estimater på, er å la estimeringen følge samme iterative arbeidsform som man gjør for arkitektur, utvikling og testing (Royce 1998). I modellens første fase – inception - er artefaktene fremdeles veldig uferdige og estimeringsstrategien blir primært Top-Down baserte ekspertestimerer. I de påfølgende fasene – elaboration og construction – vil estimeringen kunne ivaretas av ekspertvurderinger med BottomUp strategi. Uformelle spesialistgrupper eller av Prosjektleder med spesialisert har karakteristika som er godt egnet for bruk i prosjekter basert på Unified Process.

XP er en utviklingsmodell som preges av at kravene til produktet skal utformes og tilpasses underveis i prosessen. Et tidlig totalestimat strider i utgangspunktet mot denne tilnærmingen til systemutvikling, der man ønsker en utviklingsprosess med utviklere som selv estimerer de enkelte oppgavene innen den iterasjon man er inne i (Beck and Fowler 2001). Utover en horisont på de nærmeste 2-6 ukene har man meget uklare formeningar om hva som skal utvikles eller hvor mye ressurser det vil kreve. I forhold til estimeringsmetodene mener jeg at ”Ustrukturerte grupper” er et naturlig valg i de omgivelsene utviklingsarbeidet her foretas innen.

### **Referanser til publikasjoner der dette omtales spesielt :**

Fairley så for seg en utvikling innen fagfeltet i retning av at hver type utviklingsmodell fikk sin estimeringsmetode (Fairley 1992). Vi ser at de nylanserte utviklingsmodellene i sterkere grad enn før gir føringer for estimeringsarbeidet, men det lanseres ikke egne spesifikke estimeringsmetoder (Royce 1998; Kruchten 2000; Beck and Fowler 2001).

De modellbaserte metoders sterke knytning til fossefallsmodellen og de utfordringer dette gir i forhold til endringene i tilnærmingen til systemutvikling er diskutert. En angrepsvinkel har vært å tilpasse de eksisterende modellbaserte metodene til også å fungere innen mer moderne systemutviklingsmodeller (Boehm and Sullivan 1999). Hvorvidt man lykkes med å tilpasse seg de nye omgivelsene foreligger det foreløpig lite materiale på.

### **Råd :**

Ta konsekvensen av at estimering er en støtteaktivitet, og velg en estimeringsmetode som er i takt med og ikke i konflikt med den systemutviklingsmodell prosjektet følger.

## **Kategori: Tilgjengelighet på informasjon**

### **5.3.12. Data fra tidligere gjennomførte prosjekter**

#### **Beskrivelse:**

Historiske data fra beslektede gjennomførte prosjekter er viktige innen all estimering. Som det er påpekt har alle estimeringsmetodene nytte av å støtte seg på historiske data fra den aktuelle kontekst. Ekspertvurderingene anvender dette på en uformell måte, mens modellbaserte metoder nyttegjør seg dataene til å kalibrere modellene. Det er likevel innen analogibasert estimering at historiske data har den mest direkte betydningen.

Den sterke utviklingstakten innen teknologi og anvendelsesområder, gjør at tilgangen på og relevansen av historiske data er svært varierende i ulike kontekster. Det er heller ingen sterk tradisjon for å registrere faktisk forbruk av ressurser innen de ulike delene av et systemutviklingsforløp. Dette har konsekvenser for metodevalget ettersom metodene i ulik grad forutsetter tilgang på historiske data. Et viktig aspekt rundt disse dataene er kvaliteten. Man vil relativt enkelt kunne innføre rutiner for å registrere timeforbruk, men det vil i mange tilfeller ligge et sett forklaringer bak hvorfor timeforbruket ble slik, noe som er vanskelig å få inn i prosjektdatabaser.

#### **Konsekvenser for vurderingen av metoder:**

Som det er påpekt i gjennomgangen, er analogibaserte estimeringsmetoder helt avhengig av at det foreligger historiske data med relevans i forhold til den forestående utviklingsoppgave. I kontekster der man ikke sitter med formaliserte historiske data er disse metodene helt uaktuelle. De modellbaserte metodene trenger historikk for å kunne kalibreres til organisasjonskonteksten. Data-drevne metoder er helt avhengig av historiske data for å utforme selve estimeringsmodellen, mens sammensatte metodene som COBRA er mer fleksible med hensyn til hvilke historiske data som trengs.

Erfaringer fra tidligere prosjekter er en av mange informasjonskilder for ekspertene. En stor styrke ved ekspertvurderingene er deres fleksibilitet til å forholde seg til det informasjonsgrunnlaget man har mulighet til å fremskaffe i en gitt kontekst. Dermed blir også disse metodene minst rammet av mangelfull historikk, da den uformelle anvendelsen ikke gir absolutte krav. En av metodenes styrker er å estimere oppgaver man ikke har gjennomført tidligere, og det historiske datagrunnlaget ikke finnes. For Top-Down-baserte ekspertvurderinger er det likevel påpekt et strekt behov for å ha eksperter med kjennskap til et historisk prosjekt med svært sammenfallende karakteristika for at man skal kunne oppnå en god treffsikkerhet. Har man ikke dette må den mer tidkrevende Bottom-Up strategien anvendes for å oppnå treffsikre estimer.

#### **Referanser til publikasjoner der dette omtales spesielt :**

Også for dette kriteriet er gjennomgangen i kapittel 4 sentral i forhold til hvilke publikasjoner dette temaet er diskutert. Forhold rundt analogiers sårbarhet, informasjonens rolle for modellkalibrering og ekspertvurderingenes fleksibilitet er tidligere referert under vurderingen av metodene.



Den generelt mangelfulle tradisjon for å registrere data om historisk forbruk er belyst i en studie som fant at kun halvparten av organisasjonene registrerte data om forbruk av innsats (Heemstra 1992).

I situasjoner der man har et større antall historiske prosjekter å basere seg på og samtidig har stabile teknologiske omgivelser har analogibaserte metoder vist seg å være svært egnet (Hill, Thomas et al. 2000).

Top-Down strategien er sterkt avhengig av at ekspert(e) har et sammenfallende historisk prosjekt å bygge sine estimater på (Jørgensen 2003).

### **Råd**

Siden alle estimeringsmetoder har stor nytte av historiske datagrunnlag, bør etablering av rutiner for registrering av forbrukt utviklingsinnsats prioritere.

### **5.3.13. Spesifikasjonsnivå**

#### **Beskrivelse:**

Flere av de tidligere kriteriene har berørt ulike sider ved informasjonstilgjengelighet når man skal velge estimeringsmetode. Her belyses betydningen av detaljeringsgraden på informasjonen. For å predikere utviklingsinnsats er det alltid en fordel med mest mulig detaljert informasjon om oppgaven. Siden estimering er en støtteaktivitet, er man underlagt det detaljnivå man har på spesifikasjonen på det tidspunkt estimatene må gis.

#### **Konsekvenser for vurderingen av metoder:**

For ekspertvurderinger har spesifikasjonen i første rekke innflytelse på om man kan benytte seg av Bottom-Up strategi. Er kravene ikke spesifisert og aktivitetene ikke delt opp, må man følge Top-Down strategien. I slike situasjoner er "Prosjektleder med spesialister" uaktuell da denne kun er aktuell ved Bottom-Up.

Modellbaserte metoder vil gi en langt bedre kvalitet på estimatene når spesifikasjonene er detaljerte. Da kommer modellenes analytiske styrker til anvendelse. Istedenfor at sentral input om størrelsen på den fremtidige applikasjonen må baseres på ekspertuttalelser, kan man senere i utviklingsforløpet med større sikkerhet predikere størrelsen på basis av informasjon i den detaljerte kravspesifikasjonen.

Analogibaserte metoder har en klar styrke i at de ikke fordrer noen detaljerte spesifikasjon på oppgaven. Man bygger på at den spesifikasjonen som ble gjort i det utvalgte analogiprojektet er representativ for den nye løsningen. Eneste krav er at et sett overordnede karakteristika som skal danne grunnlaget for å identifisere en analogi kan angis.

#### **Referanser til publikasjoner der dette omtales spesielt :**

Jeg henviser her til kapittel 4, da informasjonsgrunnlag ble grundig belyst ved gjennomgangen av metodene.

### Råd

Utsett estimeringen så lenge dette er forsvarlig, da en mer detaljert spesifikasjon alltid vil være gunstig for estimeringen.

### 5.4. Status for rammeverket

Jeg anser mitt rammeverk som etablert nå etter at jeg har gitt en grundig gjennomgang av hvert enkelt valgkriterium. Strategien Strukturert Vurdering er blitt fulgt opp, ved at jeg har beskrevet metodealternativene og valgkriteriene. Ved å gi en dekkende kategorisering av metodealternativene og gjennomgå hver enkelt metodens sterke og svake sider i kapittel 4, ble de lagt et grunnlag som man kunne vurdere metodene ut fra. I dette kapitlet har jeg så beskrevet de tretten valgkriteriene jeg mener er nødvendig for å ivareta konteksten. Jeg har her vist hvordan man innen hvert kriterium må trekke inn metodenes styrker og svakheter for å kunne vurdere relevansen av den enkelte metode i den gitte kontekst.

Målet om å lansere et velbegrunnet forslag til hvordan man kan øke bevisstheten ved valg av metode skulle dermed være nådd. Kapitlene 3, 4 og 5 utgjør således fundamentet i rammeverket. I det neste kapitlet vil jeg gjøre et forsøk på å antyde hvordan disse byggestenene er tenkt anvendt som en helhet av de som sitter med ansvaret for å velge metode.

## 6. ANVENDELSE AV RAMMEVERKET

### 6.1. Utkast til anvendelsesveiledning

De foregående kapitlene har hatt en akademisk vinkling ved at det har vært et mål å gi velbegrunnede vurderinger av hvilke metoder man har å velge blant, og hvilke kriterier som best sikrer at konteksten gis den nødvendige innflytelse på valget av estimeringsmetode. I dette kapitlet ser jeg rammeverket fra et mer pragmatisk perspektiv, og vurderer hvordan dette kan utgjøre en helhet som kan anvendes som et hjelpemiddel for de som i praksis skal velge estimeringsmetode i et konkret prosjekt. Jeg vil understreke at dette kapitlet i stor grad bærer preg av å være et tidlig utkast.

Det er klart at det vil være en fordel for de som skal bruke rammeverket å ha inngående kjennskap til hvordan dette er blitt utformet. Man må likevel se dette også fra en praktisk synsvinkel. I den virkelige verden er estimering en støtteaktivitet og de ansvarlige for metodevalget har ofte meget begrenset med tid de kan avsette til å foreta valget. Ofte vil dette være prosjektlederen som har svært mange ansvarsområder å ivareta i en hektisk oppstartsfase på et prosjekt. Videre er det også et mål å gjøre rammeverket lettforståelig og ikke minst raskt å bruke. Formålet i dette kapitlet blir å formidle de grove trekk av hvordan den grundige gjennomgangen som gjort med hensyn til vurdering av alternative metoder og vurderingskriterier kan bli et godt hjelpemiddel i disse omgivelsene.

En forutsetning for å kunne anvende rammeverket på en fornuftig måte er at essensen i innholdet i kapitlene som beskriver metodenes styrker og svakheter (Kap. 4.4, 4.5 og 4.6), samt detaljbeskrivelsen av valgkriteriene (Kap.5.3) er kjent. Ved noen av kriteriene er det gitt konkrete råd om metodevalg, mens det andre steder kun er gitt råd som veileder den ansvarlige for metodevalget. Den tette knytningen mellom valgkriteriene og metodenes sterke og svake sider er som jeg flere steder har påpekt selve essensen i rammeverket.

En konsentrert versjon av disse kapitlene rettet mot de som i praksis skal anvende rammeverket vil være en gunstig videreutvikling. Dette ligger utenfor den fokus jeg har i mitt arbeid, og jeg trekker derfor bare frem hovedtrekk her. Understående tabell oppsummerer de styrker og svakheter jeg fant de ved ulike metodene i kapittel 4. Denne kan benyttes som en huskeliste ved bruk av rammeverket, men fritar altså ikke den ansvarlige fra selv å sette seg bedre inn de ulike metodene.

Metode	Styrker	Svakheter
Ekspertvurderinger generelt	Fleksibilitet Evne til å takle spesielle forhold	Mangelfull dokumentasjon og repeterbarhet
Individuell ekspert	Lite ressurskrevende Tillit	Skjevhetstendenser Glemmer viktige forhold Krav til breddekunnskap
Prosjektleder med Spesialister	Reell ekspertise gjør jobben	Glemmer aktiviteter Sub-optimalisering
Matematisk aggregering	Minsker skjevhetsproblematikken	Ressurskrevende Krav om tilgjengelighet
Strukturert gruppe	Redusere skjevhetstendens Motiverende	Store krav til ressurser og tilgjengelighet
Ustrukturert gruppe	Ivaretar ulike perspektiver Reduserer skjevhetstendens	Ressurskrevende
Analogibaserte metoder	Lett å bruke og gir stor tillit Egnet ved stabile forhold Tidlig tidspunkt Automatisering	Avhengig av data fra gjennomførte prosjekter Svake på usikkerhet Ustabilitet
Modellbaserte metoder generelt	Automatisering	Avhengig av nøyaktig informasjon
Proprietære metoder	Lite avhengig av estimeringsekspertise Bred anvendelse	Ikke innblikk i estimatberegning
Ikke-proprietære metoder	Lite avhengig av ekspertise estimeringsekspertise Ivaretar kompleksitet og størrelse Er dokumenterte	Lite egnet til tidlige estimerer og ved iterative utviklingsmodeller Avhengig av kalibrering
Data-drevne metoder	Estimeringsmodellen tilpasses kontekst	Krever teoretisk kompetanse
Sammensatte metoder	Kobler styrkene i to ulike tilnærminger	Høy igangkjøringskostnad Krever både informasjon og ekspertise

Tabell 6.1. Styrker og svakheter ved metodene – fra kapittel 4.

Fra kapittel 5 kan de rådene jeg utledet ved bearbeidingen av hvert kriterium utgjøre nyttige, men ikke tilstrekkelige antydninger med hensyn til bruk av rammeverket. :

1. Følg de metodevalg som eventuelt er angitt i standarder og rutiner, men gi alltid en vurdering av erfaringene i etterkant av prosjektet
2. Utsett estimeringen så lenge som mulig, da alle estimeringsmetodene nyttiggjør seg at informasjonsgrunnlaget blir bedre.
3. Formuler et mest mulig presist motiv, og se hvilken tendens dette gir i forhold til metodevalg.

4. La ikke treffsikkerheten være det dominerende valgkriteriet.
5. Avklar tidspunkter og ajourholdsfrekvensen samtidig da disse kriteriene ofte stiller ulike krav til metodene
6. Velg aldri en estimeringsmetode som fordrer større estimeringsressurser enn det du er innstilt på å bruke.
7. For de estimeringsansvarlige er opplæring innen grunnprinsippene i de alternative metodene viktigere enn kursing i detaljer innen enkeltmetoder.
8. Velg estimeringsmetode først etter å ha vurdert den reelle tilgjengeligheten på ekspertise i den gitte prosjektkontekst.
9. Bruk metoder du tidligere har hatt suksess med, og eksperter som tidligere har gitt gode estimater.
10. Velg den enkleste estimeringsmetoden som kan mestre kompleksiteten og størrelsen på oppgaven.
11. Ta konsekvensen av at estimering er en støtteaktivitet, og velg en estimeringsmetode som er i takt med og ikke i konflikt med den systemutviklingsmodell prosjektet følger.
12. Siden alle estimeringsmetoder har stor nytte av historiske datagrunnlag, bør etablering av rutiner for registrering av forbrukt utviklingsinnsats prioritere.
13. Utsett estimeringen så lenge dette er forsvarlig, da en mer detaljert spesifikasjon alltid vil være gunstig for estimeringen.

Etter at å ha gjort seg kjent med alternativene og kriteriene i rammeverket, foreslår jeg at prosjektlederen benytter understående matrise ved sin gjennomgang av det enkelte valgkriterium. Vedkommende må analyserer forholdene rundt hvert kriterium i den kontekst man har i prosjektet. Metoder som diskvalifiseres på grunnlag av de lokale forholdene gis et minustegn (-) i matrisen. De metoder som har styrker som gjør dem spesielt egnet i den gitte kontekst gis et pluss tegn i matrisen (+). I noen tilfeller er en konkret metode eneste mulige alternativ, noe som for eksempel kan gjøre seg gjeldende innen "Eksterne krav". Det markeres med et stjernetegn (\*). Dette gir et forenklet men like fullt anvendelig helhetsinntrykk.

Alternative metoder Valg-kriterier	Individuell Ekspert	Prosjektleder med spesialister	Matematisk aggregering	Strukturert gruppe-prosess	Ustrukturert gruppe-prosess	Analogi-basert	Proprietær Metode	Ikke-proprietær metode	Data-dreven metode	Sammen-satt metode
Eksterne krav										
Utsette eller kutte ut										
Motiv										
Treffsikkerhet										
Tidspunkt										
Avsatte estimerings-ressurser										
Teoretisk kompetanse										
Tilgjengelighet på eksperter										
Tidligere erfaringer med estimerings-metoder										
Størrelse på produkt og utviklings-gruppe										
System-utviklings-modell										
Data fra gjennomførte prosjekter										
Spesifikasjons-nivå										

Tabell 6.2. Matrise for metodevalg.

Etter å ha gått gjennom alle kriteriene, bør man trekke inn de tverrgående kriteriene jeg har lansert innen Ekspertvurderinger. Man bør kommentere hvordan disse påvirker og eventuelt endrer de vurderinger som ble gjort ved den første gjennomgangen. Disse kriteriene er:

- Top-Down kontra Bottom-Up strategi
- Bruk av sjekklister
- Bruk av maler
- Ekspertise behov (Domene, Teknologi, Estimering)
- Objektive kontra subjektive eksperter

Anvendelse av rammeverket skal føre til et bevisst metodevalg, men som man ser av dette utkastet til veiledning, gir det ingen automatisering av valgprosessen. Rammeverket gir i første rekke informasjon til prosjektleder/valgansvarlig om hvilke føringer hvert av de tretten kriteriene isolert gir med hensyn til metodevalg.

På bakgrunn av den ”profilen” utfylling av matrisen gir i form av +/-/\* ved de alternative metodene, må den ansvarlige selv gjøre en subjektiv totalvurdering. På dette grunnlag velges estimeringsmetode for prosjektet. Gevinsten av å benytte rammeverket

er at valgprosessen blir preget av en god struktur som sikrer en helhetsvurdering. Valgprosessen er egnet for repetisjon i senere prosjekter, men de faglige vurderingene vil fortsatt måtte foretas av brukeren.

### **6.2. Valg av metode på organisasjonsnivå og prosjektnivå**

Når man skal velge estimeringsmetode er det interessant å vurdere om man kan standardisere på en estimeringsmetode innen hele virksomheten. På den måten kan man oppnå en enhetlig måte å håndtere denne støtteaktiviteten på. Dette vil gi positive effekter i form at man etter hvert får etablert en kultur for hvordan man estimerer og kan sette opp rutiner for aktiviteten. Man vil bygge opp kompetanse som er gjenbrukbar på tvers av prosjekter, og det blir etter hvert en bevissthet rundt hva man kan forvente at ligger i grunn for estimatene som prosjektlederne presenterer.

Ved å velge metode organisasjonsnivå er det mulig å ivareta organisasjonskonteksten på en god måte. Det bør da gjennomføres en valgprosess som følger den anvendelsen av rammeverket som jeg har skissert over, men fokus må da være på fellestrekk fra den organisasjonsmessige konteksten, og ikke nede på prosjektnivå. Dette vil være en gunstig tilnærming i organisasjoner som utfører svært homogene utviklingsoppdrag, og der gevinstene som diskuteres over ved å samordne og benytte en estimeringsmetode på organisasjonsnivå er større enn den fleksibiliteten man taper innen det enkelte prosjekt.

Har man derimot et mangfold i ulike typer prosjekter, der man gjennomfører utviklingsaktivitetene i ulike kontekster, vil et metodevalg på organisasjonsnivå føre til at man nede i det enkelte prosjekt må anvende en estimeringsmetode som er lite egnet i den kontekst man har i prosjektet. Her vil et prosjektspesifikt valg av estimeringsmetode være bedre. Ved å anvende rammeverket får man et godt beslutningsgrunnlag for å velge estimeringsmetode uten at dette skal være spesielt ressurskrevende. Forutsetningene for å lage mer treffsikre estimerer med et reelt eierskap blant de ansvarlige blir bedre. Dette fordrer at man har en stor grad av omstillingsevne til å ta i bruk ulike metoder i ulike prosjekter. Rammeverket er utformet slik at man kan anvende det i begge de foreslåtte sammenhenger.

## 7. KONKLUSJON OG FORSLAG TIL VIDERE ARBEID

### 7.1. Konklusjon

I dette hovedfagsarbeidet har jeg utformet et rammeverk for å ivareta organisasjons- og prosjektkontekst ved valg av estimeringsmetode i systemutviklingsprosjekter. Selv om det er totaliteten i rammeverket som utgjør det samlede bidrag, vil jeg vil trekke frem følgende fire konklusjoner og trekk fra dette arbeidet som de viktigste:

- Jeg har vist at det er relevant å anvende tilnærminger hentet fra prediksjonsteori for å øke bevisstheten rundt valg av estimeringsmetode innen systemutvikling
- Det er argumentert for og foreslått en ny metodekategorisering som oppveier den skjevhet eksisterende inndelinger har mot modellbaserte metoder
- Evalueringskriterier for å sikre ivaretagelse av kontekst er lansert og diskutert i forhold til metodenes styrker og svakheter
- I egenskap av å vise sammenhengen mellom alternativene og vurderingskriteriene er rammeverket blitt et lovende hjelpemiddel for økt bevisstheten rundt metodevalg.

Gjennom litteraturstudiet fant jeg at valg av estimeringsmetode for systemutviklingsprosjekter i liten grad har vært et tema innen forskningen. I den grad det har vært diskutert er det med ensidig fokus på modellbaserte estimeringsmetoder. I mangelen på relevant litteratur innen eget fagfelt, har jeg vurdert publikasjoner som omhandler tilnærminger til valg av prediksjonsmetode innen feltet *Forecasting* (Makridakis, Wheelwright et al. 1983; Armstrong 2001). Min konklusjon er at de prinsipielle trekkene ved et metodevalg er de samme. Jeg har vist relevansen ved å utforme et rammeverk med utgangspunkt i en valgstrategi og tilnærminger til metodevalg herfra. Ut fra dette har jeg har utformet et robust rammeverk tilpasset omgivelsene innen systemutvikling.

Anvendelsen av *Strukturert Vurdering* som valgstrategi gjorde at jeg avdekket en mangelfull kategorisering av estimeringsmetodene innen systemutvikling. Til tross for at ekspertvurderinger er den overlegent mest utbredte tilnærmingen i systemutviklingsprosjekter, viste litteraturstudien at eksisterende metodekategoriseringer i for stor grad bærer preg av at forskningen har vært konsentrert rundt modellbaserte metoder. I mitt arbeid har jeg derfor hatt et omfattende materiale å bygge på innen modellbaserte metoder, mens jeg i langt større grad har måttet basere meg på enkeltpublikasjoner, inspirasjon fra prediksjonslitteratur og egne vurderinger når det gjelder rammeverkets dekning av ekspertvurderinger. Jeg har ved å systematisere tilnærminger som er lasert fra ulike forskningsmiljøer, kunnet lansere til forslag til en mer dekkende og balansert kategorisering av estimeringsmetodene vi har innen systemutvikling.



Litteraturstudiet ga innblikk i et stort antall publikasjoner som på ulike måter har påpekt styrker og svakheter i de enkelte estimeringsmetodene. Den akkumulerte forskningen har dermed avdekket mange interessante trekk ved metodene. Dette har likevel ikke vært brukt systematisk i sammenheng med valg av estimeringsmetode. Gjennom mitt arbeid har jeg vist at koblingen mellom metodenes styrker/svakheter og karakteristika i et prosjekts omgivelser er et nøkkelområde for å bli i stand til å velge en egnet estimeringsmetode. Mitt arbeid her er på ingen måte komplett, men det viser en del sentrale trekk i hvordan man bør legge opp prosessen for å velge estimeringsmetode.

For å ivareta konteksten er det etter min vurdering nødvendig å gå bort fra tradisjonen med å måle kvaliteten på estimeringsmetodene ut fra ett kriterium – treffsikkerheten. Ivaretagelse av kontekst stiller store krav til å se problemstillingen fra mange perspektiv. Også her er min konklusjon at vi har mye å lære fra andre fagfelt, og det er lansert tretten ulike kriterier som prosjektleder eller estimeringsansvarlig bør vurdere i sin avveining mellom metodene.

Utfordringen er selvsagt å gjennomgå tretten valgkriterier innen de rammer ressursene tillater. Man må unngå å benytte mitt rammeverk på en rigid måte der man er ute etter eksakte svar. Dette vil kreve for mye ressurser. **Budskapet er isteden at den ansvarlige, i egenskap av å kjenne den aktuelle prosjektkontekst og med støtte i rammeverket, i løpet av minutter skal gi en helhetlig og kvalifisert vurdering av hvilken metode som er best egnet i den gitte kontekst.** Fortsatt vil valgprosessen være preget av subjektive vurderinger, men disse vil være satt bedre i system.

Når man ser på grunnstrukturen i rammeverket virker innholdet i stor grad selvsagt og innlysende. På mange måter uttrykker det "sunn fornuft" satt i system. Dette er et trekk som jeg finner både viktig og interessant. Jeg mener at mitt arbeid har vist at rammeverket baserer seg på anerkjente forskningspublikasjoner. Når dette likevel resulterer i at så vel estimeringsmetodene som vurderingskriteriene virker selvsagte, ser jeg dette som en bekreftelse på at det mye å hente ved å systematisere den akkumulerte kunnskapen vi allerede har på området. Etter min oppfatning er det derfor ikke noe som tyder på at rammeverket i for stor grad baserer seg på ren intuisjon. Det andre interessante trekket er at dette kan slå positivt ut med hensyn til å oppnå aksept for å benytte rammeverket i reelle prosjekter. De påtenkte brukerne vil enklere kunne ta rammeverket i bruk ved at det er lett å forstå.

### 7.2. Forslag til videre arbeid

Før rammeverket er klart til anvendelse er det behov for å kvalitetssikre det og prøve det ut i kontrollerte omgivelser. Som nevnt i kapittel 1.4. er det endelige målet at rammeverket skal kunne tas i bruk. For mitt hovedfagsarbeid var derimot målsettingen å komme med et utkast som forskere kan gi tilbakemeldinger på, og som spesielt interesserte prosjektledere kan gå gjennom og prøve ut i kontrollerte omgivelser. Mine forslag til videreføring av arbeidet er :

### **Publisere en kortversjon av arbeidet etter at det er kvalitetssikret særlig med hensyn til referansegrunnlaget**

Referansegrunnlaget for de ulike metodenes styrker og svakheter, samt referansene som er knyttet til de enkelte valgkriterier er et område der det klart er potensiale for forbedringer. Dette er et viktig felt som jeg har lagt ned mye ressurser i å kartlegge. Det er likevel for omfattende til at jeg gjennom mitt hovedfagsarbeid har kunnet gi en dekkende gjennomgang. Min fortolkning av de vurderinger andre forskere har gjort bør kvalitetssikres. Det finnes åpenbart også relevante publikasjoner jeg ikke har klart å finne. En gjennomgang av arbeidet av erfarne forskere ville slik sett være nyttig for å luke ut feiltolkninger og samtidig gi bedre støtte for påstander med et mangelfullt referansegrunnlag. Etter en slik gjennomgang vil en publisering av arbeidet kunne skape debatt om temaet, og føre til konstruktive tilbakemeldinger fra forskningsmiljøet.

### **Videreføre kapittel 6 til å bli en ”praktikerens håndbok”**

En som vil teste ut denne tilnærmingen i praksis ønsker ikke å lese gjennom et stort antall sider med faglige vurderinger for hvordan rammeverket er blitt utformet. Det trengs derfor en kortfattet innføring og praktisk orientert veiledning. Kapittel 6 er et gunstig utgangspunkt å bygge videre på her. Veiledningen må utvides til å gi en lettfattelig fremstilling av de mest relevante delene fra kapittel 4 og 5.

### **Kontrollert utprøving**

En mindre gruppe spesielt interesserte prosjektledere bør på et egnet utvalg av prosjekter prøve ut rammeverket. Ved å legge opp dette på en vitenskapelig måte, vil en slik utprøving kunne gi nyttige innspill på hvordan rammeverket med sine alternative metoder og vurderingskriterier fungerer i praksis.

---

## REFERANSER

- Abdel-Hamid, T. K. and S. E. Madnik (1983). "The dynamics of software project scheduling." Communications of the ACM **26**(5): 340-346.
- Ambler, S. (1999). More Process Patterns - Delivering Large-Scale Systems Using Object Technology. Cambridge University Press.
- Anda B, Angelvik E, et al. (2002). "Improving Estimation Process by Applying Use Case Models." Lecture Notes in Computer Science **2559/2002**: 383-397.
- Angelis, L. and I. Stamelos (2000). "A simulation tool for efficient analogy based cost estimation." Journal of Empirical Software Engineering **5**(1): 35-68.
- Armstrong, J. S., Ed. (2001). Principles of forecasting: A handbook for researchers and practitioners. International series in operations research & management scienc. Boston, Kluwer Academic Publishers.
- Armstrong, J. S. (2001). Selecting forecasting methods. Principles of forecasting: A handbook for researchers and practitioners. J. S. Armstrong. Boston, Kluwer Academic Publishers: 365-386.
- Beck, K. and M. Fowler (2001). Planning EXtreme Programming. Boston, Addison Wesley.
- Boehm, B., C. Abts, et al. (2000). Software cost estimation with Cocomo II. New Jersey, Prentice-Hall.
- Boehm, B. and K. Sullivan (1999). "Software economics: Status and prospects." Information and Software Technology **41**(14): 937-946.
- Boehm, B. W. (1981). Software engineering economics. New Jersey, Prentice-Hall.
- Boehm, B. W. (1984). "Software engineering economics." IEEE Transactions on Software Engineering **10**(1): 4-21.
- Boehm, B. W., B. Clark, et al. (1995). "Cost models for future software life cycle processes: COCOMO 2.0." Annals of Software Engineering **1**: 57-94.
- Boehm, B. W. and R. W. Wolverton (1980). "Software Cost Modeling: Some Lessons Learned." The journal of Systems and Software: 195-201.
- Briand, L. C., K. El Emam, et al. (1998). COBRA: A hybrid method for software cost estimation, benchmarking, and risk assessment. International Conference on Software Engineering, Kyoto, Japan, IEEE Comput. Soc., Los Alamitos, USA.
- Briand, L. C. and I. Wiczorek (2002). Resource estimation in software engineering. Encyclopedia of software engineering. J. J. Marcinak. New York, John Wiley & Sons.
- Chambers, J. C., S. K. Mullick, et al. (1971). "How to choose the right forecasting technique." Harvard Business Review **49**(4): 45-74.
- Chatzoglou, P. D. and L. A. Macaulay (1996). "A review of existing models for project planning and estimation and the need for a new approach." International Journal of Project Management **14**(3): 173-183.
- Connolly, T. and D. Dean (1997). "Decomposed versus holistic estimates of effort required for software writing tasks." Management Science **43**(7): 1029-1045.
- Conte, S. D., H. E. Dunsmore, et al. (1986). Software engineering metrics and models. Menlo Park, California, Benjamin Cummings.
- Coombs, P. (2003). IT Project Estimation - A Practical Guide to the Costing of Software. Cambridge University Press.
- DeMarco, T. (1982). Controlling software projects. New York, Yourdon Press.
- Dolado, J. J. (2001). "On the problem of the software cost function." Information and Software Technology **43**(1): 61-72.
- Fairley, D. (2002). "Making Accurate Estimates." IEEE Software **19**(6): 61-63.
- Fairley, R. E. (1992). Recent advances in software estimation techniques. International Conference on Software Engineering, Melbourne, Australia, ACM, New York, NY, USA.
- Fenton, N. and S. L. Pfleeger (1996). Software Metrics, A Rigorous & Practical Approach. Cambridge, PWS Publishing Company.
- Ferens, D. V. (1999). "The conundrum of software estimation models." IEEE Aerospace and Electronic Systems Magazine **14**(3): 23-29.

- 
- Ferrell, W. R. (1985). Combining Individual Judgments, Chapter 6 in Wright G. ed Behavioral Decision Theory. New York, Plenum Press.
- Finnie, G. R., G. E. Wittig, et al. (1997). "A comparison of software effort estimation techniques: Using function points with neural networks, case-based reasoning and regression models." Journal of Systems and Software **39**(3): 281-289.
- Gray, A., S. MacDonnell, et al. (1999). Factors systematically associated with errors in subjective estimates of software development effort: the stability of expert judgment. Sixth International Software Metrics Symposium, IEEE Comput. Soc, Los Alamitos, CA, USA.
- Halvorsen, K. (1993). Å forske på samfunnet - en innføring i samfunnsvitenskapelig metode, Bedriftsøkonomisk forlag.
- Hammond, K. R. (1996). Human judgement and social policy: Irreducible uncertainty, inevitable error, unavoidable injustice. New York, Oxford University Press.
- Heemstra, F. J. (1992). "Software cost estimation." Information and Software Technology **34**(10): 627-639.
- Hill, J., L. C. Thomas, et al. (2000). "Experts' estimates of task durations in software development projects." International Journal of Project Management **18**(1): 13-21.
- Hughes, R. T. (1996). "Expert judgement as an estimating method." Information and Software Technology **38**(2): 67-75.
- Höst, M. and C. Wohlin (1997). "A subjective effort estimation experiment." Information and Software Technology **39**(11): 755-762.
- Höst, M. and C. Wohlin (1998). An experimental study of individual subjective effort estimations and combinations of the estimates. International Conference on Software Engineering, Kyoto, Japan, IEEE Comput. Soc, Los Alamitos, CA, USA.
- Jacobsen, D. I. (2000). Hvordan gjennomføre undersøkelser ? Innføring i samfunnsvitenskapelige metoder. Kristiansand, Høyskoleforlaget.
- Jørgensen, M. (1995). "Experience with the accuracy of software maintenance task effort prediction models." IEEE Transactions on Software Engineering **21**(8): 674-681.
- Jørgensen, M. (2002). "A Review of Studies on Expert Estimation of Software Development Effort." Submitted to Journal of Systems and Software.
- Jørgensen, M. (2003). "A Preliminary Checklist for Software Cost Management." QSIC 2003.
- Jørgensen, M. (2003). "Top-Down and Bottom-Up Expert Estimation of Software Development Effort." Submitted to Journal of Information and Software Technology 2004.
- Jørgensen, M. and K. Moløkken (2002). Combination of Software Development Effort Prediction Intervals: Why, When and How? To appear in proceedings of: IEEE Conference on Software Engineering and Knowledge Engineering, Italy.
- Jørgensen, M. and D. I. K. Sjøberg (2001). "Impact of effort estimates on software project work." Information and Software Technology **43**(15): 939-948.
- Jørgensen, M. and D. I. K. Sjøberg (2001). "Software process improvement and human judgement heuristics." Scandinavian Journal of Information Systems **13**: 99-121.
- Jørgensen, M., D. I. K. Sjøberg, et al. (2000). "The prediction ability of experienced software maintainers." IEEE Computer Society: 93-100.
- Jørgensen, M. and K. H. Teigen (2002). Uncertainty Intervals versus Interval Uncertainty: An Alternative Method for Eliciting Effort Prediction Intervals in Software Development Projects. To appear in proceedings of: International conference on Project Management (ProMAC), Singapore.
- Jørgensen, M., K. H. Teigen, et al. (2002). "Better Sure than Safe? Overconfidence in Judgment Based Software Development Effort Prediction Intervals." Submitted to Journal of Systems and Software.
- Kemerer, C. F. (1987). "An empirical validation of software cost estimation models." Communications of the ACM **30**(5): 416-429.
- Kitchenham, B., S. L. Pfleeger, et al. (2002). "A case study of maintenance estimation accuracy." To appear in: Journal of Systems and Software.
- Kruchten, P. (2000). The Rational Unified Process - an Introduction Second Edition. Boston, Addison Wesley.
- Lederer, A. L., R. Mirani, et al. (1990). "Information system cost estimating: a management perspective." MIS Quarterly **14**(2): 159-176.
- Lederer, A. L. and J. Prasad (1992). "Nine management guidelines for better cost estimating." Communications of the ACM **35**(2): 51-59.

- 
- Lederer, A. L. and J. Prasad (1998). "A causal model for software cost estimating error." IEEE Transactions on Software Engineering **24**(2): 137-148.
- Londeix, B. (1987). Cost estimation for software development, Addison-Wesley.
- Londeix, B. (1995). "Deploying realistic estimation (field situation analysis)." Information and Software Technology **37**(12): 655-670.
- Mahmoud, E., R. DeRoeck, et al. (1992). "Bridging the gap between theory and practice in forecasting." International Journal of Forecasting **8**(2): 251-267.
- Makridakis, S., S. C. Wheelwright, et al. (1983). Forecasting : methods and applications. New York, John Wiley & Son.
- Maxwell, K. D., L. van Wassenhove, et al. (1996). "Software Development Productivity of European Space, Military, and Industrial Applications." IEEE Transactions on Software Engineering **22**(10): 706 ..
- McCulla, P. (1989). "The estimating process (IT project management)." International Journal of Project Management **7**(1): 36-38.
- Meyer, M. A. and J. M. Booker (1991). Eliciting and analyzing expert judgment: A practical guide. Philadelphia, Pennsylvania, SIAM.
- Moløkken, K. (2002). Expert estimation of Web-development effort: Individual biases and group processes (Master Thesis). Department of Informatics, University of Oslo.
- Moløkken, K. and M. Jørgensen (2002). "Expert estimation of the effort of web-development projects: Why are software professionals in technical roles more optimistic than those in non-technical roles." unpublished.
- Moløkken, K. and M. Jørgensen (2003). "Software Effort Estimation: Unstructured Group Discussion and a Method to Reduce Individual Biases." PPIG bidrag 2003.
- Mukhopadhyay, T., S. S. Vicinanza, et al. (1992). "Examining the feasibility of a case-based reasoning model for software effort estimation." MIS Quarterly **16**(2): 155-171.
- Navlakha, J. K. (1990). "Choosing a software cost estimation model for your organization: A case study." Information and Management **18**(5): 255-261.
- Park, R. E. (1996). "A manager's checklist for validating software cost and schedule estimates." American Programmer **9**(6): 30-35.
- Reifer, D. (2000). "Web Development: Estimating Quick-to-Market Software." IEEE Software(6): 57-64.
- Royce, W. (1998). Software Project Management - A Unified Framework. Reading, Addison Wesley.
- Ruhe, M., R. Jeffery, et al. (2003). "Cost Estimation for Web Applications." Proceedings, 25th ICSE, Portland 2003: 285-294.
- Schooff, R. M. and Y. Y. Haimes (1999). "Dynamic multistage software estimation." IEEE Transactions on Systems, man, and cybernetics **29**(2): 272-284.
- Shanteau, J. (1987). "Psychological Characteristics of Expert Decision Makers." NATO ASI Series F35: 289-304.
- Shepperd, M. and C. Schofield (1997). "Estimating software project effort using analogies." IEEE Transactions on Software Engineering **23**(11): 736-743.
- Shepperd, M., C. Shofield, et al. (1996). Effort estimation using analogy. International Conference on Software Engineering, Berlin, Germany, IEEE Comput. Soc. Press, Los Alamitos, CA, USA.
- Taff, L. M., J. W. Borchering, et al. (1991). "Estimateings: development estimates and a front-end process for a large project." IEEE Transactions on Software Engineering **17**(8): 839-849.
- Vicinanza, S. S., T. Mukhopadhyay, et al. (1991). "Software effort estimation: An exploratory study of expert performance." Information systems research **2**(4): 243-262.
- Walkerden, F. and D. R. Jeffery (1997). "Software cost estimation: A review of models, process, and practice." Advances in Computers **44**: 59-125.
- Walkerden, F. and R. Jeffery (1999). "An empirical study of analogy-based software effort estimation." Journal of Empirical Software Engineering **4**(2): 135-158.
- Winklhofer, H., A. Diamantopoulos, et al. (1996). "Forecasting practice: a review of the empirical literature and an agenda for future research." International Journal of Forecasting **12**(2): 193-221.

---

## VEDLEGG 1 : ENGELSK VERSJON AV SAMMENDRAG

I dette vedlegget ligger den engelske versjonen av sammendraget. Denne ble levert elektronisk i forbindelse med innleveringen av oppgaven.

### **Selection of Software Cost Estimation Methods – A Framework Regarding Project – and Organisation Context**

Abstract :

Project managers and organisation management want early and accurate predictions on the effort needed to develop software. A number of alternative estimation methods have been proposed by researchers to support this. The alternative methods all tend to have drawbacks and advantages, making them suitable in different environments. Empirical results on cost estimation accuracy of software projects suggest that the selection of software cost estimation method should be context-dependent.

Selecting the most appropriate among the variety of existing estimation methods regarding organisational and project context is important. This topic has almost been neglected in software cost estimation research. Developing and evolving model-based estimation techniques have dominated research. In this thesis I present a framework on software cost estimation method selection. The work is based on a study of published literature on software cost estimation and general prediction theory.

The framework is based on a selection strategy called Structured Judgment. Here you evaluate the available estimation methods on a set of selection criteria. First I suggest a categorization of existing methods to sketch the alternatives available. Existing categorization have been evaluated, but found to be biased towards model-based methods. Therefore a new categorization is presented, elaborating different types of expert judgment as well.

Then a number of selection criteria are presented. Research has been dominated by publications regarding accuracy as the only criteria. In my framework thirteen selection criteria are suggested to be able to consider context in an appropriate way. I propose selection criteria divided into the categories: 1) Process constraints, 2) Purpose and use, 3) Estimation capability, 4) Project characteristics and 5) Information availability.

This framework has been developed considering the fact that estimation is a supportive activity with limited resources available to make a selection among alternative methods.