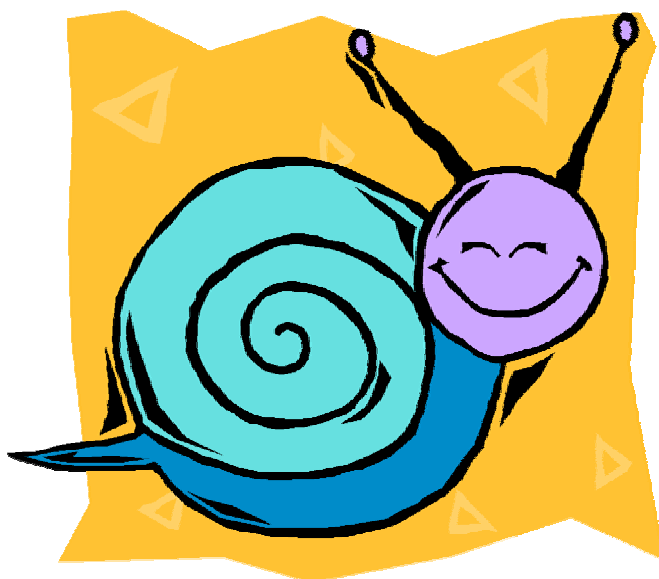


Benny Jansson, Erik Lindberg, Malin Persson

Snabbsnigelmetoden

En iterativ metod för utvärdering av simuleringsprogram



TOTALFÖRSVARETS FORSKNING SINSTITUT

Försvarsanalys
172 90 Stockholm

FOI-R--0728--SE

Februari 2003

ISSN 1650-1942

Metodrapport

Benny Jansson, Erik Lindberg, Malin Persson

Snabbsnigelmetoden

En iterativ metod för utvärdering av simuleringsprogram

Utgivare Totalförsvarets Forskningsinstitut - FOI Försvarsanalys 172 90 Stockholm	Rapportnummer, ISRN FOI-R--0728--SE	Klassificering Metodrapport
	Forskningsområde 2. Operationsanalys, modellering och simulering	
	Månad, år Februari 2003	Projektnummer E 1416
	Verksamhetsgren 5. Uppdragsfinansierad verksamhet	
	Delområde 22 Metod och utredningsstöd	
Författare/redaktör Benny Jansson Erik Lindberg Malin Persson	Projektledare Karin Mossberg Sonnek	
	Godkänd av Elisabeth André Turlind	
	Uppdragsgivare/kundbeteckning Försvarsmakten	
	Tekniskt och/eller vetenskapligt ansvarig	
Rapportens titel Snabbsnigelmetoden - En iterativ metod för utvärdering av simuleringsprogram		
Sammanfattning (högst 200 ord) Det finns idag gott om metoder för att utveckla programvara, men få för utvärdering av befintliga simuleringsprogram. Syftet med denna rapport är att presentera en metod som ska kunna ge analytiker stöd vid utvärdering av simuleringsprogram som används, eller eventuellt kommer att användas, vid FOI Försvarsanalys. Snabbsnigelmetoden är inriktad på översiktliga utredningar, vilket innebär att metoden hellre ska vara snabb och ge ett översiktligt resultat snarare än att ge en fullständig analys. Genom en syntes av litteraturstudier och intervjuer med erfarna analytiker har ett antal slutsatser genererats, av vilka främst följande två ligger till grund för metodens utformning: 1. Det är lämpligt att inleda utvärderingen med att undersöka i vilken utsträckning programmet behandlar rätt frågor innan resurser förbrukas på att testa om svaren är korrekta. 2. Det är den potentiella användaren som bäst kan bedöma om programmet besvarar rätt frågor. Följaktligen ägnas en stor del av metoden åt att hitta och formulera relevanta frågeställningar samt att undersöka i vilken utsträckning dessa behandlas av programmet, medan endast en mindre del ägnas åt verifiering och validering av resultaten från simuleringen. Arbets sättet benämndes snabbsnigelmetoden eftersom den iterativa arbetsprocessen kan liknas vid formen av ett snigelskal och tonvikten ligger på snabbhet.		
Nyckelord Utvärdering, simulering, modell, metod, värdering, validering		
Övriga bibliografiska uppgifter	Språk Svenska	
ISSN 1650-1942	Antal sidor: 44 s.	
Distribution enligt missiv	Pris: Enligt prislista	

Issuing organization FOI – Swedish Defence Research Agency Defence Analysis SE-172 90 Stockholm	Report number, ISRN FOI-R--0728--SE	Report type Methodology report
	Research area code 2. Operational Research, Modelling and Simulation	
	Month year February 2003	Project no. E 1416
	Customers code 5. Commissioned Research	
	Sub area code 22 Operational Analysis and Support	
Author/s (editor/s) Benny Jansson Erik Lindberg Malin Persson	Project manager Karin Mossberg Sonnek	
	Approved by Elisabeth André Turlind	
	Sponsoring agency The armed forces	
	Scientifically and technically responsible	
Report title (In translation) The Quick-coil method		
Abstract (not more than 200 words) Today there are several existing methods available for developing computer simulation programs. However, there are no such methods in place for evaluating existing programs at FOI's Department of Defence Analysis. The aim of this report is to present a method which will give support to analysts in evaluating simulation programs that are, or will be, used at FOI. The Quick-coil method focuses on general surveys, which means that the method is designed to provide the analyst with a quick synopsis rather than an exhaustive analysis. As a result of a comprehensive study, several conclusions were reached regarding the requirements for such a method. The two most important of which were used as the basis for the design of the Quick-coil method: 1. It is critical to ensure that the program addresses the right questions before resources are spent on verifying the resultant answers. 2. The potential program user is the individual best suited to determine whether the program asks correct questions. Consequently, the Quick-coil method is focused on defining the relevant questions and testing whether the program answers these questions. Less effort is spent on validation and verification of simulation results. The evaluation process is called the Quick-coil method, since the iterative nature of the process can be illustrated as a coil, and that the method is designed to provide timely results.		
Keywords Evaluation, simulation, model, method, validation		
Further bibliographic information	Language Swedish	
ISSN 1650-1942	Pages 44 p.	
Price acc. to pricelist		

Innehåll

1	INLEDNING.....	6
1.1	BAKGRUND.....	6
1.2	SYFTE.....	6
1.3	PROBLEMFORMULERING	6
1.4	AVGRÄNSNINGAR OCH DEFINITIONER.....	7
1.5	GENOMFÖRANDE.....	10
2	VAL MELLAN BEHOVS- OCH VERKTYGSFOKUSERAD METOD.....	11
2.1	ALLMÄNT	11
2.2	KONSEKVENSBESKRIVNING	11
3	SNABBSNIGELMETODEN.....	13
3.1	ANTAGANDEN	13
3.2	METODENS FEM STEG.....	13
	<i>Steg 1: Översiktlig programbeskrivning.....</i>	<i>16</i>
	<i>Steg 2: Alternativa simuleringsprogram</i>	<i>20</i>
	<i>Steg 3: Undersökning av behov och förmåga.....</i>	<i>20</i>
	<i>Steg 4: Fördjupad undersökning.....</i>	<i>23</i>
	<i>Steg 5: Dokumentation.....</i>	<i>27</i>
4	TILLÄMPNING AV METODEN.....	28
4.1	ALLMÄNT	28
4.2	NIMROD 5W	28
4.3	TREAMM VERSION 1.0	29
5	SAMMANFATTNING	31
6	REFERENSER.....	32
BILAGA 1.	CHECKLISTA	35
BILAGA 2.	V&V-METODIK INOM MOS.....	39
BILAGA 3.	EGENUTVECKLING KONTRA INKÖP AV MODELLER	41

1 Inledning

1.1 Bakgrund

Modellering och simulering (MoS) har stor betydelse för FOI Försvarsanalys och det är betydelsefullt att säkerställa kvaliteten på de modeller som används. Processen för kvalitetssäkring av modeller brukar kallas VV&A vilket står för Verifiering, Validering och Ackreditering (André *et al* 1999).

Det finns idag gott om metoder för att utveckla programvara men det finns ingen, för FOI Försvarsanalys, fastställd metod för utvärdering av befintlig simuleringsprogramvara. En sådan metod skulle kunna:

- Hjälpa programanvändare att snabbare fokusera på de viktigaste frågorna som bör besvaras när ett simuleringsprogram utvärderas.
- Säkerställa att kvaliteten på de studier av simuleringsprogram som genomförs blir bättre och jämnare.

1.2 Syfte

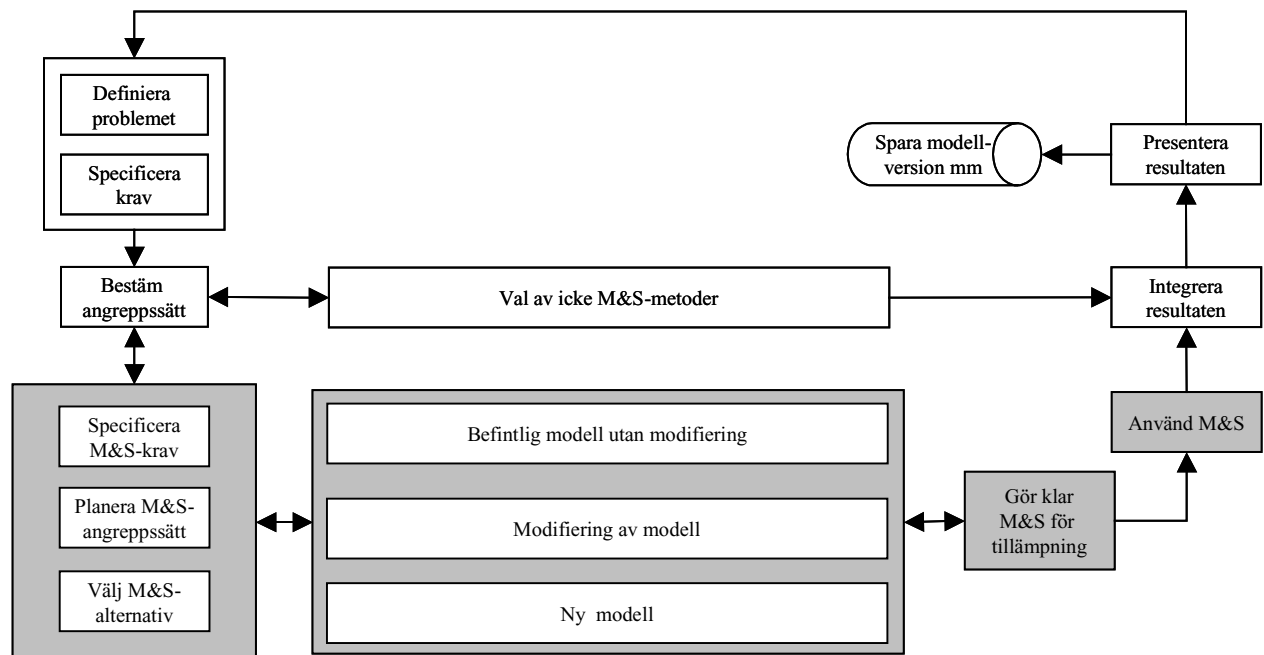
Syftet med denna rapport är att presentera en metod som ska kunna ge analytiker stöd vid utvärdering av simuleringsprogram som används, eller eventuellt kommer att användas, vid avdelningen för Försvarsanalys. Rapporten riktar sig främst till anställda vid FOI Försvarsanalys med mindre erfarenhet av simuleringsprogram och utvärderingsarbete.

1.3 Problemformulering

Att undersöka användbarheten för befintliga simuleringsprogram är egentligen ett "bakvänt" arbetssätt då det utgår från verktyget, och inte från behovet. Ett effektivare arbetssätt är att utgå ifrån behovet och förut-sättningslöst söka efter ett passande verktyg (simuleringsprogram). Då detta "rättvända" arbetssätt tillämpas kan metodiken för VV&A som beskrivs av André *et al* (1999) användas, se *Figur 1*.

En orsak till att det "bakvända" arbetssättet tillämpas kan exempelvis vara att en analytiker kommit i kontakt med ett intressant simuleringsprogram som denne skulle vilja veta mer om, utan att ha resurser för att kunna göra en fullständig utredning. Författarnas avsikt har därför varit att skapa en metod som kan lösa uppgiften så snabbt och effektivt som möjligt. Ett sätt att

beskriva metoden är att den tillämpar ett *verktysfokuserat* angreppssätt istället för en *behovsbaserad* problemlösning.



Figur 1. Förenklad figur av VV&A in the Problem Solving Process (DMSO). Figuren beskriver en iterativ process med början i det övre vänstra hörnet, "Definiera problemet". Fullständig figur och beskrivning återfinns i FOI rapport "Verifiering, validering och ackreditering av simuleringsmodeller", André *et al* (1999)

1.4 Avgränsningar och definitioner

En grundläggande avsikt är att metoden främst ska vara snabb och ge ett översiktligt resultat. Det är därför lämpligt att börja med att undersöka om programmet kan besvara rätt frågor innan resurser förbrukas på att testa om svaren är korrekta. På grund av detta ägnas endast en mindre del av metoden åt verifiering och validering av resultaten från simuleringen. En större del ägnas åt att ge modellförståelse genom att undersöka vilka frågor som besvaras av programmet samt om svaren på dessa är intressanta för användaren.

Under intervjuer med analytiker på FOI Försvarsanalys identifierades ett problem avseende generell metodanvändbarhet kontra avgränsat tillämpningsområde. Är metoden för generell riskerar den att bli för urvattnad för att ge tillräckligt stöd. Är den för problemspecifik begränsas tillämpningsområdet för mycket för att vara praktiskt användbart. Författarna försöker därför i denna rapport avgränsa tillämpningsområdet på ett sådant sätt att det varken

blir för generellt eller för snävt. För att göra denna avgränsning måste dock begreppen modell och simulering först definieras:

- *Modell.* En modell är en approximation av en viss företeelse. Den är en förenkling av företeelsen genom att den fokuserar på vissa egenskaper. Mer specifikt betraktas här modeller som kan representeras matematiskt eller logiskt.
- *Simulering.* En simulering är ett experiment med en modell. Experimentet avser att följa och förstå modellens beteende och orsakssammanhang under ett tidsförlopp.

Enligt Hughes (1997) klassificeras vanligtvis de olika typerna av simuleringar enligt nedanstående:

- 1) *Live simulation.* A simulation in which real people operate real systems.
- 2) *Virtual simulation.* A simulation in which real people operate simulated systems.
- 3) *Constructive Model or Simulation.* Models and simulations that involve simulated people operating simulated systems. Real people stimulate (make inputs to) such simulations but are not involved further in determining the outcomes.

Utifrån denna definition avgränsas tillämpningsområdet till att endast gälla för den tredje typen av simulering.

För att ytterligare avgränsa tillämpningsområdet behandlas endast simuleringsprogram som är kopplade till den väpnade striden inom områdena väpnat angrepp, territoriell integritet och internationella insatser. Hughes (1997) anser att sju olika kategorier kan definieras för denna typ av tillämpning:

- 1) *Battle Planning.* Preparation for wartime operations, based on friendly and enemy orders of battle, the existing strategic or tactical environment, and specific missions or military objectives. Battle planning models encompass duels, engagements, battles, and campaigns, including combat-oriented operations away from the battlefield.

- 2) *Wartime Operations*. The conduct of war. Distinguished from battle planning by available, current, wartime data and known, immediate military objectives.
- 3) *Weapon Procurement*. The selection from among competing weapon systems or characteristics for procurement decisions.
- 4) *Force Sizing*. The decision of how many weapon systems of which types to (1) operate, (2) support, and (3) procure in the future, either in the defense establishment as a whole or in a major component such as the Army or the nuclear weapons arsenal.
- 5) *Human Resource Planning*. The design and operation of manpower, personnel, training, and assignment systems.
- 6) *Logistic Planning*. The design and operation of all manners of military logistic support.
- 7) *National Policy Analysis*. Supramilitary actions that influence or are influenced by military considerations such as arms treaties or subsidies of commercial transportation.

Dessa kategorier borde kunna kombineras med den indelning som används av André *et al.* (1999) för att beskriva Försvarmaktens användning av modell- och simuleringsstöd. Dessa är:

- Studier och analys
- Utbildning och träning
- Systemutveckling och materielanskaffning
- Ledning och planering
- Forskning och utveckling

Hittills har metoden bara testats på två av dessa kategorier:

- Battle Planning – Studier och analys (*Nimrod 5W*, Lindberg 2002)
- Weapon Procurement – Systemutveckling och materielanskaffning (*TREMM Version 1.0*, Jansson 2002)

Författarna har inte haft möjlighet att bedöma huruvida metoden är användbar för de övriga kategorierna. Förhoppningsvis kan detta visas genom att framtida användare kommer att pröva metodens användbarhet på andra typer av programvara.

1.5 Genomförande

Idén till snabbsnigelmetoden föddes med avstamp i två likartat formulerade uppgifter vilka båda syftade till att undersöka användbarheten hos var sitt simuleringsprogram: *Nimrod 5W* som simulerar flygstrid av traditionell karaktär på operativ nivå respektive *TREMM Version 1.0* som simulerar minsvep mot flera minor. Det framkom att inga vedertagna metoder fanns för hur denna typ av granskning bör genomföras, varför arbetet med snabbsnigelmetoden initierades.

För att skapa en utvärderingsmetod för programvara, gjorde författarna ett antal antaganden angående rimliga angreppssätt. Dessa antaganden verifierades/falsifierades genom intervjuer med personer på FOI vilka alla hade olika erfarenheter av modelleringsrelaterade uppgifter. Intervjuerna bidrog till nya diskussioner om infallsvinklar till problemet och modifiering av de ursprungliga antagandena. Detta förfarande upprepades ett antal gånger innan en utvärderingsmetod formulerades. Arbetet inkluderade också sökning och genomläsning av litteratur i såväl tryckt form, som på nätet, i jakt på liknande studier och metoder av intresse.

Tillämpning av metoden redovisas kortfattat i kapitel 4 samt i sin helhet i två separata rapporter: *TREMM Version 1.0* (Jansson 2002) och *Nimrod 5W* (Lindberg 2002).

2 Val mellan behovs- och verktygsfokuserad metod

2.1 Allmänt

Vid programutvärdering är det lockande att börja med att försöka validera de resultat som simuleringsprogrammet levererar. Författarna anser att detta direkta angreppssätt är felaktigt. Istället bör arbetet inledas med en analys av vilka följder (konsekvenser) som användning av simuleringsprogrammet kan leda till. Orsaken till detta är att det är klokt att låta storleksordningen på dessa följder påverka valet av utvärderingsmetod. Denna slutsats formuleras också av André *et al* (1999):

”Hur stor VV&A-insats som är tillräcklig eller nödvändig för en speciell tillämpning är i första hand beroende av tillämpningens art och syfte. Det är viktigt att en ordentlig riskanalys genomförs för att rätt kunna bedöma hur omfattande verifierings- och valideringsarbete som är nödvändigt.”

2.2 Konsekvensbeskrivning

Enligt ovanstående resonemang är det lämpligt att börja med att strukturera och beskriva de indirekta konsekvenser som resultatet av simuleringarna kan leda till, d.v.s. vilken typ av beslut programmet kan ge underlag för. Exempel på användningsområden och deras indirekta konsekvenser kan vara:

- Indata för studie vilken kan leda till *anskaffning/avveckling av system*, exempelvis ett nytt luftförsvarssystem. I detta fall kan det vara viktigt att grovt försöka bedöma de kostnader som beslutet kommer att kunna påverka.
- Indata för studie vilken kan medföra stor påverkan genom *organisationsförändringar*, exempelvis framtagning av nya metoder för strategiska transporter vilket kräver en ny logistikfunktion. I detta fall bör utvärderaren försöka uppskatta om felaktiga resultat från programmet kan leda till onödigt stora kostnader.
- Indata för studie som kan påverka förändringar avseende *taktiskt uppträdande*. Om denna typ av beslut fattas på felaktiga grunder kan det medföra risk för person- och/eller materielskador.
- Användning av programvaran som *utbildningsredskap* vilket kommer att påverka tänkande och operativt eller taktiskt uppträdande hos beslutsfattare/operatörer. På samma sätt som ovan kan felaktiga resultat från programmet medföra risk för person- och/eller materielskador.

Konsekvenserna är svåra att bedöma exakt, men det räcker här med en grov uppskattning då huvudsyftet är att utvärderaren tänkt igenom vilken påverkan användning av programmet kan ge. För mer djupgående resonemang avseende risk och konsekvenser vid modellanvändning hänvisas läsaren till rapporten *Ett riskperspektiv på verifiering, validering och ackreditering av simuleringsmodeller* (Bergsten 2002).

Om konsekvensbeskrivningen indikerar att det finns stor risk för person- och/eller materielskador, alternativt onödigt höga kostnader, är det klokt att utgå ifrån behovet istället för verktyget. *The Defense Modeling and Simulation Office* (DMSO) har utarbetat en metod för hur dessa problem angrips. Denna metod beskrivs utförligt av André *et al* (1999).

Om konsekvenserna däremot kan anses vara av mindre allvarlig karaktär är det troligt att en verktygsfokuserad metod är användbar. De problem som behandlas på FOI Försvarsanalys är oftast av långsiktig och övergripande karaktär med många kontrollstationer före implementering, vilket motiverar användande av en verktygsfokuserad metod.

3 Snabbsnigelmetoden

3.1 Antaganden

Snabbsnigelmetoden beskriver arbetet stegvis och är konstruerad för att användas då utvärderingen utgår från programverktyget istället för behovet. Angreppssättet är framställt för utvärderingsuppgifter i mindre skala med krav på systematisk dokumentation. Metoden bygger på ett antal antaganden som formulerats utifrån analys av intervjuer och litteraturstudier. De viktigaste av dessa antaganden är:

- 1) Det går inte att avgöra om ett simuleringsverktyg är användbart utan att först ta reda på vilka frågor som användaren har behov av att behandla. Utvärderingsarbetet bör därför inledas med en undersökning om i vilken utsträckning programmet behandlar rätt frågor. Först därefter är det meningsfullt att undersöka om svaren är rätt.
- 2) Intervjuer med potentiella användare av programmet är ett effektivt sätt att hitta och beskriva frågeställningar som det finns behov av att behandla.
- 3) För att skapa förståelse hos potentiella användare avseende vilken typ av frågeställningar som söks, bör intervjuer inledas med en översiktlig beskrivning av programmet.
- 4) Att bedöma användbarheten av det givna programmet utan att fundera över alternativa verktyg innebär risk för suboptimering. För att minska denna risk bör därför en översiktlig inventering av olika alternativa simuleringsprogram göras. Ambitionsnivån på denna inventering beror givetvis på anslagna kostnadsramar.

3.2 Metodens fem steg

Utifrån ovanstående antaganden formuleras metoden i fem steg:

Steg 1: *Översiktlig programbeskrivning.* Först görs en översiktlig programbeskrivning där programmet schematiskt sammanfattas och klassificeras avseende modelltyp, syfte och användningsområde. Detta ska ge översiktlig förståelse för hur programmet fungerar samt vad det är konstruerat att användas till. Dessa kunskaper används sedan som underlag för undersökningen i steg 3.

Steg 2: *Alternativa simuleringsprogram.* När utvärderaren i steg 1 bildat sig en uppfattning om programmet skapas underlag för jämförelse genom att söka

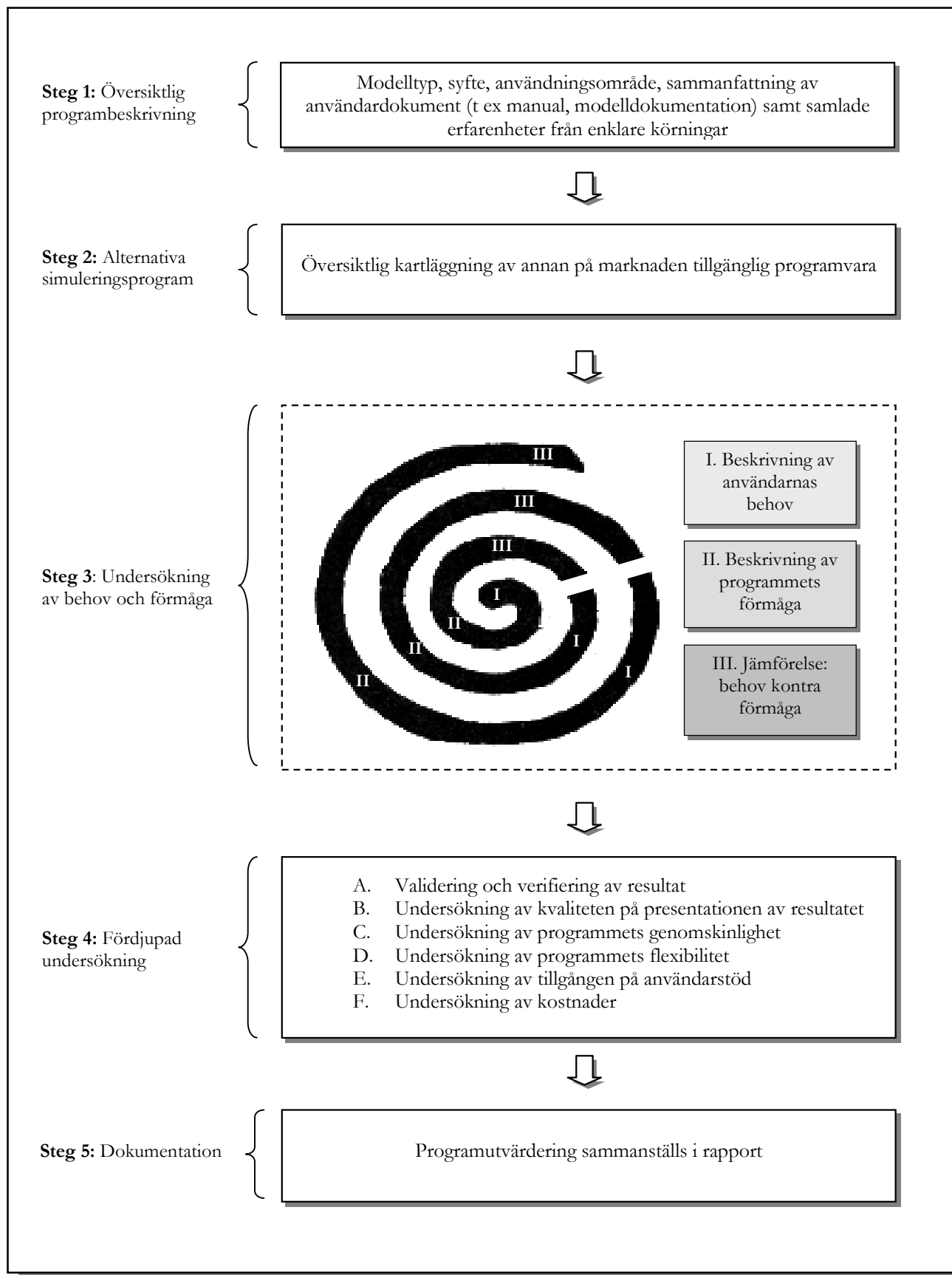
andra liknande program och kortfattat beskriva dessa, samt om möjligt klassificera dessa enligt samma standard som använts i steg 1. Detta material är dels viktigt för att genom jämförelse kunna värdera det aktuella programmens förmåga, dels för att försöka ge en bild av alternativa sätt att lösa uppgiften. Resultatet från steg 2 ger också utvärderaren övergripande kunskap om vad som tidigare gjorts inom det aktuella problemområdet.

Steg 3: *Undersökning av behov och förmåga.* I detta steg presenteras programmet kortfattat för en potentiell användare varefter denne får beskriva ett antal möjliga behov. Därefter undersöks programmens förmåga att uppfylla dessa behov. Vid behov av noggrannare utvärderingsresultat kan arbetet upprepas. Tidigare erhållna resultat används då som grund för den upprepade processen. Denna iterativa metodik kan illustreras som en spiral/snigelskal, se *Figur 2*.

Steg 4: *Fördjupad undersökning.* Kan programmet troligtvis uppfylla en potentiell användares behov är det lämpligt att undersöka programmens trovärdighet, användbarhet och flexibilitet. Sist görs en bedömning av de kostnader som kan förknippas med bruket av programmet.

Steg 5: *Dokumentation.* I detta avslutande steg sammanställs och dokumenteras erfarenheterna och resultatet av utvärderingsarbetet.

För att ge en uppfattning om hur stora resurser arbetet kommer att ta i anspråk i respektive steg bör utvärderingens ändamål och ambitionsnivå bestämmas i ett tidigt skede.



Figur 2. Schematisk bild över snabbsnigelprocessen

Steg 1: Översiktlig programbeskrivning

Avsikten med detta steg är att sammanställa en översiktlig programbeskrivning som ska användas som underlag i steg 3. Det är lämpligt att börja med att beskriva information som kan hämtas ur manualer eller motsvarande användardokument. Denna information bör därefter kompletteras med erfarenheter från enkla programkörningar. För att skapa en tydlig bild av programmets egenskaper är det viktigt att beskriva 1) *Programmets ursprungliga syfte och användningsområde* och 2) *Typ av modell* samt 3) *Karaktären på indata och utdata*.

Programmets ursprungliga syfte och användningsområde

Det finns en mängd olika sätt att definiera en simuleringsmodells syfte. Som tidigare nämnts beskriver André *et al* (1999) försvarsmaktens användning av modell- och simuleringsstöd med en indelning i fem kategorier: *Studier och analys*, *Utbildning och träning*, *Systemutveckling och materielanskaffning*, *Ledning och planering* och *Forskning och utveckling*.

Enligt Hughes (1997) är det lämpligt att definiera modellen enligt *Models by scope or scale* (modellens omfattning eller nivå):

- *Campaign or theater-level models* beskriver strider på operativ nivå och över större områden.
- *Battle or multi-unit engagement models* beskriver sammandrabbningar mellan större styrkor.
- *Single engagement models* beskriver sammandrabbningar mellan enskilda enheter, t ex u-båt mot u-båt, d.v.s. duellsituationer.
- *Phenomenological models* är de mest grundläggande och beskriver beteendet hos enskilda fysiska enheter (inklusive mänskliga).

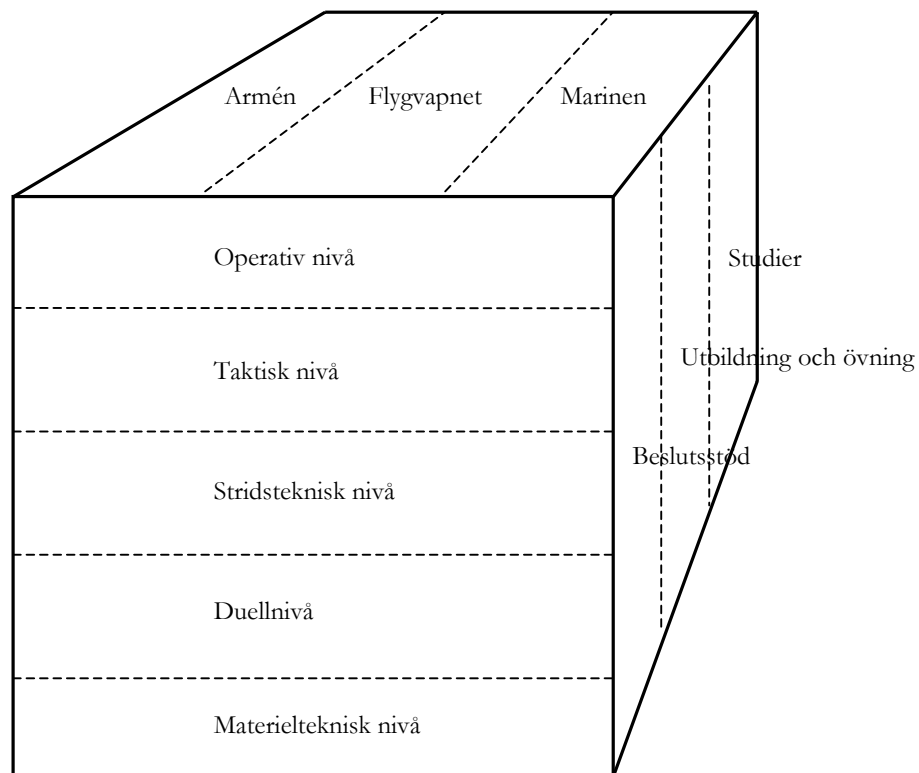
På svenska kan detta indelas på följande sätt:

- *Operativ nivå*. Försvarsgrensgemensam (joint) verksamhet. Den operativa nivån ska se till att de militära ansträngningarna bär mot ett gemensamt mål som har undernivåerna *luftoperativ*, *markoperativ* och *sjöoperativ* verksamhet. Storleksordningen på ingående enheter är exempelvis brigad. Tidsförloppet är veckor eller dagar. Utsträckningen i rummet är av milostorlek, t. ex. norra Sverige.
- *Taktisk nivå*. Storleksordningen på ingående enheter är exempelvis bataljon i markstriden eller ”många mot många”-situationer i luft-

striden. Tidsförloppet är enstaka dygn. Utsträckningen i rummet kan vara 10 gånger 10 mil i markstriden eller södra Östersjön i sjöstriden.

- *Stridsteknisk nivå*. Ingående enheters storlek är exempelvis kompani eller rotar av flygplan eller fartyg och tidsförloppet är timmar eller minuter.
- *Duellnivå*. Ingående enheter är exempelvis enstaka stridsvagnar, flygplan eller fartyg. Tidsförloppet är minuter.
- *Materielteknisk nivå*. Fokus kan t.ex. vara på effekten av olika typer av stridsdelar eller sensorer. Tidsförloppet är minuter eller sekunder.

Sammanställs ovanstående parametrar med hänsyn till om simuleringsmodellen tagits fram för armén, marinen eller flygvapnet ges nedanstående figur¹:



Figur 3. Olika dimensioner vid beskrivning av simuleringsmodeller. Omarbetad för denna rapport efter USA M&S Masterplan (1995). (Odar, *IT som stöd för strategiska beslut*, 2002, s 7)

¹ Den traditionella indelningen mellan mark, sjö och luft är användbar för de flesta nuvarande simuleringsprogram men bör för framtida program eventuellt ersättas med en indelning som fokuserar på det nätverksbaserade försvaret.

Klassificering av modelltyp

Nedan ges kortare förklaringar till olika vanligt förekommande begrepp vid modellklassificering.

Analytisk modell/ Simuleringsmodell

En analytisk modell anger direkt genom formler hur olika indata påverkar slutresultatet. Till ett enkelt system kan på detta sätt ett exakt resultat erhållas. Användaren får en bra överblick över, och förståelse för, sammanhangen mellan indata och resultat men inte vad som hänt på vägen. Simulering innebär att en modell prövas numeriskt för att se hur indata påverkar utdata och används då uppgiften inte kan lösas analytiskt (Law & Kelton 2000).

Statisk/Dynamisk

En statisk modell representerar ett system vid en specifik tidpunkt eller ett system där tiden inte spelar någon roll. En dynamisk modell representerar ett system som involverar tid. (Law & Kelton 2000). En terrängmodell av ett fjällandskap är ett exempel på en statisk modell, medan en modell över årstidsvariationerna i snötäcket i fjällvärlden är ett exempel på en dynamisk modell (Ulla-Stina Åström, FOI).

Deterministisk/ Stokastisk

I en deterministisk modell tas inte hänsyn till den slumpvisa variationen hos ingående storheter. För givna indata är resultatet bestämt (Law & Kelton 2000). En stokastisk modell tar hänsyn till den slumpvisa variationen. Detta innebär att parametrarna i funktionerna, ekvationerna och olikheterna som bygger upp modellen inte har bestämda värden, utan istället är angivna med en fördelningsfunktion (stokastiska). Resultatet från en modellexekvering ger antingen hela utfallsrummet (statistisk modell), eller endast ett möjligt utfall. För att i det senare fallet se spridningen i resultatet måste modellen exekveras upprepade gånger.

Kontinuerlig/ Diskret

En kontinuerlig modell är en modell vars variabler ändras kontinuerligt med tiden, medan en diskret modell har variabler som ändras momentant, d.v.s. endast vid vissa tidpunkter. För en tidskontinuerlig simulering är simuleringsmodellens interna klocka synkron, d.v.s. den stegar fram simulerad tid i jämna steg. Tidskontinuerlig simulering kallas också för *tidsstyrd simulering* (*Metodhandbok flödessimulering, d4ISy*). Den vanligaste formen av tidsdiskret simulering är *händelsestyrd simulering*, där klockan hela tiden stegas fram till tiden för nästa

händelse. Datorer med binära beräkningsmodeller är i princip alltid tidsdiskreta.

Modellexempel

En vanlig analytisk modell är Lanchesters linjära lag som beskriver en schematiserad modell av ett stridsförlopp. Denna metod behandlas inte här men beskrivs utförligt i tidskriften *FOA orienterar OM operationsanalys* (1967).

Vid simuleringar används ofta Monte Carlo-metoden. Monte Carlo-modeller beskriver händelsekedjor i kronologisk ordning. Dessa händelsekedjor innehåller ett antal alternativa vägar, vilka var och en är försedd med en sannolikhet för att just den vägen skall väljas (Law & Kelton 2000). I tidskriften *FOA orienterar OM operationsanalys* (1967) beskrivs Monte Carlo-metodiken enligt följande:

”...ett slags statistisk försöksverksamhet där man för att uppskatta lösningen på ett problem konstruerar ett hasardspel som har den egenskapen att det genomsnittliga resultatet av spelet överensstämmer med lösningen på det aktuella problemet.”

En Monte Carlo-modell kan köras en gång (enparti Monte-Carlo) eller med flera iterationer (flerparti Monte-Carlo).

Beskrivning av indata och utdata

För att öka den potentielle användarens förståelse vid presentationen av programmet är det viktigt att fastställa vilken sorts indata som behövs för att köra modellen och vilken typ av utdata programmet ger, samt på vilken form utdata ges.

Ett bra sätt att beskriva detta är att först fastställa vilka typer av objekt, fysiska enheter, programmet hanterar, t.ex. flygplan, minsvep etc. Därefter preciseras vad dessa objekt kan göra i modellen. Detta kan uttryckas som att de substantiv och verb som hanteras av programmet ska beskrivas. Det är också värdefullt att vid beskrivningen precisera modellens upplösning i tid och rum, samt om modellen är kompatibel med andra program, t.ex. ordbehandlare och databaser.

Utvärderaren bör också dokumentera huruvida programmet är HLA-kompatibelt eller inte. Förkortningen HLA står för *High Level Architecture* och är en standard för datagränssnittet för simuleringsprogram, vilket gör att olika

HLA-kompatibla simuleringsprogram kan köras tillsammans. HLA-anpassade program ger också en möjlighet till informationsutbyte med motsvarande system inom NATO. För en mer detaljerad beskrivning av HLA se hänvisning i referenslistan.

Steg 2: Alternativa simuleringsprogram

Det kan ofta vara enklare att utvärdera ett program om det finns något att jämföra med. Undersök därför om det finns andra liknande program som skulle kunna utgöra möjliga alternativ. För att fastställa hur lika eller olika dessa andra program är rekommenderas att dessa, om möjligt, klassificeras enligt ovan beskrivna klassificeringsmetod.

För att förenkla eventuella fortsatta programstudier är det viktigt att de alternativa programvarorna dokumenteras i rapporten. Eftersom de alternativa programvarorna främst ska användas som referensobjekt behöver de i övrigt inte undersökas och beskrivas lika grundligt som det ursprungliga programmet.

Steg 3: Undersökning av behov och förmåga

”Ett varv i snigeln”

Nedanstående tre faser beskriver arbetsordningen för den iterativa utvärderingsprocessen:

- I. *Beskrivning av användarnas behov.* Undersök vilka grupper som kan tänkas utgöra troliga användare, exempelvis analytiker, studiegrupper eller beslutsfattare. Definiera dessa ”troliga användares” behov genom att från intervjuer, materielplaner samt annan relevant litteratur samla in kunskaper inom aktuellt problemområde. Formulera dessa behov i frågor och frågeställningar och gör en behovsbeskrivning.

Laack *et al* (2000) påpekar att en modell bara behöver beakta de aspekter av verkligheten som är av värde för en viss applikation. Vid modellutvärdering bör således en analys av den avsedda användningen göras. Utifrån denna analys kan sedan en beskrivning göras av vilka behov som modellen måste uppfylla. Denna behovsbeskrivning bör innehålla:

- En tydlig beskrivning av det avsedda syftet med simuleringen.
- En förteckning av simuleringsbehovet, d.v.s. de fysiska enheter som måste simuleras och de funktioner de måste utföra (jmf

”substantiv och verb”) samt nödvändig grad av naturtrogenhet (degree of fidelity).

- En beskrivning av miljön som dessa enheter ska verka i, om detta är viktigt i modellen.

II. *Beskrivning av programmets förmåga.* Testa programmets förmåga mot frågeställningarna och behovsbeskrivningen. Beskrivningen av det befintliga simuleringsprogrammets förmåga bör innehålla:

- En förteckning av de fysiska enheter som ingår och de funktioner de utför (jmf ”substantiv och verb”) samt deras ”degree of fidelity”.
- En beskrivning av miljön som dessa enheter verkar i.
- En sammanfattning av antaganden och begränsningar i programkonstruktionen som kan komma att påverkar framtida simuleringar.

III. *Jämförelse: behov kontra förmåga.* Jämför behovsbeskrivningen med beskrivningen av programmets förmåga. Dokumentera vilka av de givna frågeställningarna programmet kan besvara/inte kan besvara. Om det är möjligt bör även de alternativa simuleringsprogrammets möjlighet att uppfylla behoven redovisas.

Författarna rekommenderar att samtliga tre faser i snigeln inledningsvis avverkas snabbt och översiktligt, motsvarande en förstudie. Med inhämtat material från denna inledande process beslutas sedan om ytterligare ett, nu mer utförligt, utvärderingsvarv skall genomföras. Varje varv som slutförs ger utvärderingen en noggrannare definition av möjliga framtida användningsområden och bidrar till bättre programkännedom. Den förväntade nyttan av ny kunskap måste dock vägas mot kostnaden av att gå ytterligare ett varv i utvärderingsprocessen.

Beslut

Efter ett avslutat varv i snigeln avgörs om utvärderingen bör fortsätta eller om resultatet av det senaste varvet är tillräckligt för att fylla utvärderingsarbetets syfte. Då utredningsvarvet ger resultatet att programmet kan uppfylla användarens behov helt eller delvis är det lämpligt att gå vidare till steg 4.

Uppfyller programmet inte behoven är det tillgång till tid och resurser som avgör omfattningen av den fortsatta utredningen. Antingen avslutas utredningen med slutsatsen att programvaran inte uppfyller användarens behov eller så fortsätter utredningen med en undersökning av huruvida de framtagna alternativa programvarorna kan uppfylla behoven istället.

Utvärderaren bör alltid överväga om något av de alternativa programmen som kartlagts i steg 2 kan utgöra ett bättre alternativ än det aktuella programmet, d.v.s. i större utsträckning uppfylla framtida användares behov.

Om alternativet ”avsluta utredningen” väljs kan det vara lämpligt att försöka uppskatta de risker som är förknippade med att *felaktigt* välja att *inte* använda/införskaffa programvaran. Riskerna ska vägas mot kostnaden för att köpa in programmet och för eventuella utbildnings- och underhållskostnader.

Exempel på risker kan vara:

- Att man i framtiden kan missa uppdrag eller fördröja arbetsprocessen jämfört med om man haft tillgång till programmet.
- Att man i onödan utvecklar program som redan finns.
- Att man missar att få tillgång till kunskap i kompetensnätverk som är värdefullare än de resultat som själva programmet ger.
- Att man missar viktig kunskap som finns lagrad i algoritmer eller i modellens parametrar.

Dessutom bör utvärderaren beakta möjligheten att programmet kan vara ”bra att ha” för att komplettera modellbiblioteket.

Steg 4. Fördjupad undersökning

I de fall då det är troligt att programmet uppfyller användarens behov är det lämpligt att undersöka programmets trovärdighet, användbarhet och flexibilitet samt de kostnader som är kopplade till fortsatt användning av programmet. Arbetet sker enligt följande (beskrivs mer utförligt nedan):

- Validering och verifiering av resultat.
- Undersökning av kvaliteten på presentationen av resultatet.
- Undersökning av programmets genomskinlighet.
- Undersökning av programmets flexibilitet.
- Undersökning av tillgången på användarstöd.
- Undersökning av kostnader.

I *Bilaga 3* redovisas en genomförd utredning som exempel på användbara formuleringar och synpunkter. Utredningen är inte genomförd enligt snabb-snigelmetoden men den behandlar liknande frågeställningar.

Validering och verifiering av resultat

Begreppen validering och verifiering definieras av André *et al* (1999) enligt följande:

”Med validering avses en process som avgör till vilken grad en modell är en riktig representation av verkligheten ur perspektivet av modellens avsedda användning. Med verifiering avses en process som avgör om en modellimplementation noggrant representerar utvecklarens idébeskrivning och specifikationer.”

Ett annat sätt att beskriva begreppen är att valideringen svarar på om modellen är rätt för de givna frågeställningarna och verifieringen kontrollerar om modellen är rätt byggd.

Det finns olika typer av validering, med olika noggrannhet när det gäller att avgöra trovärdigheten hos simuleringen. Validering kan ske mot andra simuleringar (”benchmarking”), mot experters kunskap (”face validation”) eller mot testdata, t.ex. resultat från övningar och fältförsök. Den sistnämnda är vad som vanligen avses när begreppet validering används. Värde av att validera mot andra simuleringar beror på hur trovärdiga dessa är. Likaså är en

acceptabel "face validation" beroende av att rätt personer utvärderar rätt data för den tänkta applikationen och att detta dokumenteras noggrant.

Vanligen anses valideringar mot testdata som de mest trovärdiga. Det finns emellertid en rad problem även med denna metod. Exempelvis kan relevanta data från strider vara kostsamma eller svåra att erhålla. Trovärdigheten hos de testdata som används måste dessutom kunna styrkas genom dokumentation av hur dessa förvärvats (Laack *et al* 2000). Olika valideringsmetoder presenteras ytterligare i *Bilaga 2*.

För snabbsnigelmetoden är antagligen "face validation" lämpligast p.g.a. att utvärderingens omfattning är begränsad. Denna kan då utföras genom att någon eller några experter bedömer om utdata från ett antal simuleringar är trovärdiga och om modellen reagerar korrekt på förändringar i indata. För de fall då valideringen visar att modellen inte är korrekt kan det bli nödvändigt att undersöka var felet ligger, d.v.s. verifiering av modellen. Denna process är dock komplicerad eftersom felkällan sällan är uppenbar. Det behöver ju inte vara fel i källkoden utan det kan lika gärna vara fel i handhavandet eller i givna indata.

Då snabbsnigelmetoden endast är tänkt för begränsade utredningar bör verifieringen helst utelämnas eller endast utföras i reducerad omfattning, eftersom simuleringsprogram ofta kräver mycket omfattande och komplexa mjukvarutester för att otvetydigt påvisa modellens riktighet (Laack *et al* 2000).

Undersökning av kvaliteten på presentationen av resultatet

En viktig del i hur trovärdigt programmet upplevs är användargränssnittets utformning. Detta är både på gott och på ont eftersom det är lätt att få förtroende för en elegant grafisk presentation även då bakomliggande algoritmer är svaga. Det är därför viktigt att försöka analysera användargränssnittet, exempelvis genom att beskriva hur programmet presenterar resultatet, vilka svar som ges samt hur dessa ges:

- Kan data matas in på ett enkelt och överskådligt sätt?
- Kan simuleringen följas grafiskt i valfri upplösning?
- Är det lätt att tillfälligt stoppa för att få ut mellandata, modifiera indata och sedan fortsätta simuleringen?
- Ger presentationen av resultatet ett trovärdigt intryck?

- Kan de utdata som användaren är intresserad av erhållas på ett snabbt och bekvämt sätt?

Som exempel på möjlighet att påverka resultatframställningen kan nämnas att simuleringsprogrammet *Nimrod 5W* inte ger användaren möjlighet att bestämma om programmet skall presentera resultaten grafiskt eller inte (Lindberg 2002). Simuleringsprogrammet *TREAMM version 1.0* tillåter däremot användaren att antingen visualisera resultatet grafiskt eller endast ge information numeriskt (Jansson 2002).

Undersökning av programmets genomskinlighet

Genomskinlighet betyder i dessa sammanhang möjligheten att förstå relationen mellan verkan och effekt i programmet. Enligt Hughes (1997) är begreppet användarspecifikt d.v.s. bedömningen av graden av genomskinlighet är beroende av användarens bakgrund (armégeneralens bedömning kommer att skilja sig från matematikerns). Genomskinligheten bör därför bedömas utifrån vem användaren är, exempelvis analytikern, studiegruppen eller beslutsfattaren. Några viktiga frågor som bör besvaras är:

- Hur överskådligt upplevs programmet?
- Är det lätt/svårt att förstå hur indata behandlas av programmet för att ge de erhållna resultaten?

Ingen modell är en exakt avbildning av verkligheten. För att bygga en modell måste alltid en mängd avgränsningar och antaganden göras. För att kunna känna tilltro till en modell och använda den på ett riktigt sätt är det viktigt att känna till dessa avgränsningar och antaganden.

Undersökning av programmets flexibilitet

En mycket viktig faktor för att beskriva ett programs egenskaper är hur väl programmet kan anpassas till nya förhållanden samt vilka möjligheter användaren ges till eget inflytande. Kan modellen anpassas till svenska förhållanden om den är utländsk, d.v.s. finns det möjlighet att göra tillägg för att beskriva svensk materiel eller svenska miljöfaktorer?

Det är även viktigt att beskriva vilken kompetens som krävs för att kunna utföra ändringarna. Måste användaren exempelvis kunna hantera ett specifikt programspråk för att modifiera programmet?

Undersökning av tillgången på stöd till användaren

För att kunna använda en modell på ett korrekt sätt är det nödvändigt att känna till hur den fungerar. Orsaken till detta är bland annat följande två faktorer som Laack *et al* (2000) pekar på:

”[S]imulations are credible only within a well-defined usage context, and only when they are properly used within that context”.

Att känna till denna “usage context” utan tillgång till omfattande och tydlig dokumentation är antagligen omöjligt. André *et al* (1999) skriver:

”Det bör observeras att man genom att acceptera att använda en ärvd modell också accepterar de inbyggda antaganden och begränsningar som finns. Att känna till dessa är näst intill omöjligt om dokumentationen av den konceptuella modellen är dålig eller helt saknas.”

Användbarheten är dock inte enbart beroende på tillgången till användardokument. Laack *et al.* anser att ett simuleringsprogramms trovärdighet och användbarhet ökar med mängden ”Simulation user support features”. Dessa är avsedda att minska möjligheterna att modellen används felaktigt. Författarna drar därför slutsatsen att dessa ”features” utgör ett viktigt kriterium vid bedömningen av användbarheten. Begreppet skulle kunna översättas med *funktioner för användarstöd*, vilka skulle kunna vara:

- Tillgång till kurser i korrekt hantering av simuleringsprogrammet och tolkning av dess utdata.
- Tillgång till omfattande dokumentation (olika manualer och användardokument). Kan användaren exempelvis utifrån manualen förstå kopplingen mellan indata och utdata?
- Tillgång till användarsupport.
- Tillgång till läsbar kod.
- Tillgång till användargrupper (user groups) som träffas regelbundet.
- Närvaron av en grundlig konfigurationshanteringsprocess som tillser att de olika versionerna av modellen hålls uppdaterade.
- Tillgång till kunniga simuleringsoperatörer på arbetsplatsen som kan se till att kunskapen om den korrekta användningen av simuleringsprogrammet inte faller i glömska med tiden.

Undersökning av kostnader

Gör en skattning av de kostnader som användningen av programvaran kan komma att medföra. Exempel på sådana kan vara:

- Anskaffningskostnader
- Driftskostnader
- Utbildningskostnader
- Kostnader för att ta fram indata
- Kostnader för att anpassa modellen till aktuella förhållanden
- Underhållskostnader för att bevara modellens aktualitet

Steg 5: Dokumentation

När användandet av snigeln bedöms vara klart sammanfattas insamlat material och presenteras på ett för utvärderingens syfte lämpligt sätt. Exempelvis kan det här vara bra att sammanfatta behovet av modifieringar vilka då kan vidarebefordras till mjukvarutillverkaren för kostnadsbedömning. I dokumentationen bör också riktlinjer för eventuellt fortsatt arbete anges.

4 Tillämpning av metoden

4.1 Allmänt

På FOI har i dagsläget två utvärderingsarbeten använt sig av snabbsnigel-metoden. Dessa presenteras kort nedan. De fullständiga arbetena presenteras i rapporterna *Nimrod 5W – Utvärdering för FoRMA/Luft och LFS 02* (Lindberg 2002) samt *TREAMM Version 1.0 – Programutvärdering med användarsynpunkter* (Jansson 2002).

4.2 NIMROD 5W

Behov av simuleringsstöd har funnits flera år, t.ex. inom ramen för FORMA/Luft. Under 2002 tillkom behov från luftförsvarsstudien (LFS 02). Ett alternativ har varit programmet NIMROD 5W, utvecklat av TIL Defence Systems Ltd Israel, vilket används vid FHS i utbildningssyfte. NIMROD var tänkt som ersättning eller komplement till det arbete som bedrivs under ANABASIS-projektet. Utvärderingen av Nimrod 5W hade därför följande huvudsyften:

- 1 Undersök om Nimrod 5W kan användas som stöd för FoRMA/Luft och LFS 02.
- 2 Undersök om Nimrod 5W kan användas för att validera ANABASIS-modellen genom att jämföra resultaten från simuleringar av liknande scenarier.

Dessutom hade utvärderingen följande alternativsyften:

- Påbörja uppbyggnaden av ett modellbibliotek som stöd för luftstridsstudier.
- Utveckla samarbetet mellan FOI och FHS.

Vid FHS används Nimrod 5W för utbildning av officerare på högre taktisk eller luftoperativ nivå. Programmet simulerar flygstrid av traditionell karaktär på operativ nivå. Enstaka duellsituationer modelleras inte explicit utan uppkomna situationer avdöms med hjälp av tabelldata. Resultatet från simuleringen erhålls i form av antal genomförda uppdrag, antal skadade mål och flygplan samt den förbrukade mängden flygbränsle och vapen.

I utvärderingen genomfördes alla stegen i snabbsnigelmetoden. FoRMA-/Luft-projektets behov av simulering var formulerat i två rapporter. Användarnas behov kunde därför extraheras ur dessa istället för ur regelrätta intervjuer. I steg 3 framkom att Nimrod 5W inte uppfyllde dessa behov. En allvarlig begränsning var att modellen inte kan behandla kryssningsrobotar.

Genom att gå ytterligare ett varv i snigeln och fokusera på uppgiftens alternativsyften visade det sig, utifrån intervju med Tore Isacson (FOI), att Nimrod 5W kunde fylla en annan typ av behov. Möjligen kan programmet användas som avdömningsverktyg vid analytiska spel eftersom det ger möjlighet att modellera krigets "friktion". Utvärderingen gick därför vidare med steg 4. Resultatet av detta steg blev att en eventuell ny version av Nimrod skulle kunna vara användbar för FOI; förutsatt att vissa delar av programmet modifieras och att dokumentationen av modellen väsentligt förbättras.

4.3 TREMM Version 1.0

FOI har tillgång till minsimuleringsprogrammet TREMM Version 1.0 (*Target Risk Estimation Against Multiple Mines*). Programmet har utvecklats av Coastal Systems Station vid Naval Surface Warfare Center i Panama City för att uppskatta effektivitet hos magnetiska minsvep och kvarvarande risk för målfartyg efter att ett område avsökts. TREMM har hittills inte använts i Sverige, varför det inom FOI inte finns erforderlig kunskap om modellen. Följaktligen initierades en programutvärdering.

TREMM är designat i syfte att utvärdera svepoperationer samt för att validera resultat från andra minsimuleringsprogram. Programmet fokuserar på en lägre taktisk/stridsteknisk nivå, då dess främsta egenskap är att producera statistiskt underlag för svepoperationer.

Utvärderingen begränsades främst av tidsmässiga skäl till att, enligt snabbsnigelmetodens numrering, belysa *steg 1* till och med *steg 3*. *Steg 4* ersattes av ett kortare avsnitt där användarsynpunkter och erfarenheter från programkörningar översiktligt presenterades. Enligt utvärderingsmetoden gjordes först ett snabbt inledande varv i snigeln. Detta skapade större förståelse för programmet och ringade in möjliga användningsområden. Därefter gjordes ytterligare ett varv vilket genererade tillräckligt med information för att presentera efterfrågat resultat. Dock rekommenderas att FOI även låter utföra *steg 4* i metoden, eftersom det är oklart om användaren kan modellera de minor, svep och fartygssignaturer som behövs för att hantera aktuella frågeställningar.

Resultatet av den genomförda utvärderingen visade att TREAMMs främsta fördel är programmets förmåga att generera statistik vid körningar med ett eller flera svep/målfartyg mot flera minor. Programmet bör i egenskap med denna förmåga vara användbart inom flera områden.

Utvärderingen visade också på att den version av TREAMM som FOI idag har tillgång till, har flera begränsningar som innebär problem för användaren. Den största begränsningen är att programmet endast behandlar minor med magnetisk påverkan samt kontaktminor. Vidare förklaras inte i manualen hur programmet räknar och användaren har svårt att få förståelse för hur specifika parametrar påverkar resultatet. Programmets källkod är emellertid tillgänglig, vilket betyder att personer som behärskar programspråket C kan skapa sig en bild av hur modellen är uppbyggd.

5 Sammanfattning

Syftet med denna rapport har varit att beskriva en metod för att utvärdera simuleringsprogram. Den tänkta målgruppen är främst analytiker med mindre erfarenhet av simuleringsprogram och liknande utredningar.

Att utgå ifrån verktyget istället för att utgå ifrån användarens behov kan tyckas något bakvänt. Eftersom det trots allt existerar ett behov av den typen av undersökningar har vår avsikt varit att skapa en metod som kan lösa uppgiften snabbt och ge ett översiktligt resultat.

Arbets sättet benämndes snabbsnigelmetoden eftersom den iterativa arbetsprocessen kan liknas vid formen av ett snigelskal och tonvikten ligger på snabbhet istället för på noggrannhet. Verifiering och validering har endast givits begränsat utrymme medan en väsentlig del av metoden ägnats åt att formulera de frågeställningar som är relevanta för användaren, samt åt att undersöka om dessa frågeställningar besvaras av programmet.

I rapporten beskrivs viktiga risker som bör beaktas vid valet av metod, samt vid beslut om huruvida programvaran ska rekommenderas eller inte. Metoden finns sammanställd i kortform i *Bilaga 1* i syfte att ge stöd åt analytikern när denne ska genomföra sin egen utvärdering.

6 Referenser

6.1 Tryckt litteratur

ANDRÉ T, BERGSTEN U, SPORRE L, 1999. *Verifiering, validering och ackreditering av simuleringsmodeller – en förstudie*, FOA--R-99-01146-202--SE, Totalförsvarets forskningsinstitut, Stockholm.

BERGSTEN U, 2002. *Ett riskperspektiv på verifiering, validering och ackreditering av simuleringsmodeller*, FOI-R--0738--SE, Totalförsvarets forskningsinstitut, Stockholm.

FOA orienterar om operationsanalys, skriftserien *Om*, nr 8, september 1967. Försvarets Forskningsanstalt.

HUGHES, W P, 1997. *Military Modeling for Decision Making*, 3rd ed., The Military Operations Research Society.

JANSSON, B, 2002. *TREAMM Version 1.0 – Programutvärdering med användarsynpunkter*, FOI-R--0577--SE, Totalförsvarets forskningsinstitut, Stockholm.

KRUTCHEN P, 1999. *The Rational Unified Process – An introduction*, 2nd ed., Addison-Wesley Longman, ISBN 0-201-60459-0

LAW, A M & KELTON, W D, 2000. *Simulation modeling and analysis*, 3rd ed. McGraw & Hill

LINDBERG E, 2002. *Nimrod 5W – Utvärdering för FoRMA/Luft och LFS 02*, Totalförsvarets forskningsinstitut, Stockholm. Rapport under arbete.

ODAR S, 2002. *IT som stöd för strategiska beslut*, Linköpings Universitet, Thesis 916, LiU-Tek-Lic-2001:53, ISBN/ISSN 9173732249

Användning av Modellering och Simulering för Flygvapenändamål (SAMS/F), steg 1, arbetsgruppens rapport nr 1, SAMS/F rapport 96:01.

6.2 Internet

DMSO, VV&A Recommended Practices Guide, *The Overall Problem Solving Process*, <http://www.msiac.dmsomil/vva/default.htm>, 2002-07-29

Laack R., Muessig P.R., Wroblenski J.J. *An Integrated Approach to Evaluating Simulation Credibility*, Summer Computer Simulation Conference 2000, www.scs.org, 2002-08-05

Gothia Science Park, Höskolan Skövde, *Projekt dAISy*, http://mechsys.ite.his.se/daisy/Metodhandbok/Metodhandbok_komplett.pdf, 2002-08-05

DMSO, HLA, *High Level Architecture*: <https://www.dmsomil/public/transition/hla/>, 2002-07-25

6.3 Intervjupersoner

FOI Försvarsanalys: Tommy André, Eva Andersson, Henrik Carlsen, Jan Frelin, Carin Kaunitz

FOI Systemteknik, Systemmodellering: Gunnar Holm, Vahid Mojtahed

Bilaga 1. Checklista

I denna bilaga ges en sammanfattning i punktform av den föreslagna metoden för utvärdering av modellerings- och simuleringsprogram.

1.1 Konsekvensbeskrivning

Bedöm de indirekta konsekvenser som användningen av programmet kan komma att leda till, det vill säga vilken typ av beslut programmet ska ge underlag för. Avgör därefter hur det fortsatta arbetet bör bedrivas. Välj metodinriktning med avseende på konsekvenser och kostnader:

- Behovsfokuserad metod (Stora konsekvenser)
- Verktýgsfokuserad metod (Mindre allvarliga konsekvenser)

Endast den verktýgsfokuserade metoden beskrivs vidare i denna bilaga. För behovsfokuserade tillvägagångssätt, se metoderna föreslagna i *Verifiering, validering och ackreditering av simuleringsmodeller*, André *et al* (1999)

1.2 Verktýgsfokuserad metod: Snabbsnigeln

1.2.1 Steg 1 Översiktlig programbeskrivning

Beskriv och klassificera programmet enligt nedanstående begrepp:

- *Programmets ursprungliga syfte och användningsområde*: Vilken beslutsnivå är programmet tänkt för (operativ, taktisk, stridsteknisk, etc.)? Är gränssnittet HLA-kompatibelt?
- *Modelltyp*: Hur behandlar programmet uppgifterna? (Analytisk-/simulering, deterministisk/stokastisk, kontinuerlig/diskret., etc.)
- *Indata/utdata*: Vilken sorts indata behövs? Vilka utdata ges och i vilken form? Vilka ”substantiv/verb” hanteras? Vilken upplösning i tid och rum simuleras (timmar, minuter, sekunder, km², m², etc.)?

Denna beskrivning ska genomföras relativt snabbt. Undvik att fastna i tidskrävande simuleringar eller att gräva ned dig i programkod. Det kan vara bättre att försöka utnyttja intervjuer med erfarna användare, text i manualerna och endast några korta programkörningar som indata.

1.2.2 Steg 2 Alternativa simuleringsprogram

Beskriv översiktligt liknande program som skulle kunna utgöra alternativ, i den mån det finns sådana. Detta material ska dels användas för att jämföra programmets förmåga, dels för att försöka ge en bild av andra sätt att lösa uppgiften. En mer detaljerad undersökning görs endast då den ursprungliga programvaran visats vara olämplig och andra alternativ söks.

1.2.3 Steg 3 Undersökning av behov och förmåga

Programmets användbarhet undersöks iterativt i tre steg:

- I. Sök troliga användares behov. Denna behovsbeskrivning bör innehålla:
 - En tydlig beskrivning av den avsedda användningen av simuleringen.
 - En förteckning av de fysiska enheter som måste simuleras och de funktioner de måste utföra (jmf ”substantiv och verb”) samt nödvändig grad av naturtrogenhet (degree of fidelity).
 - En beskrivning av miljön som dessa enheter ska verka i.
 - II. Testa programmets förmåga. Denna beskrivning av förmågan bör innehålla
 - En förteckning av de fysiska enheter som ingår och de funktioner de utför (jmf ”substantiv och verb”) samt deras ”degree of fidelity”.
 - En beskrivning av miljön som dessa enheter verkar i.
 - En sammanfattning av de antaganden och begränsningar i designen som påverkar det område i vilket simuleringen kan användas på ett trovärdigt sätt.
 - III. Bedöm hur programmets förmåga svarar mot användarbehov genom att jämföra behovsbeskrivningen med beskrivningen av programmets förmåga. Dokumentera vilka av de givna frågeställningarna programmet kan besvara/inte kan besvara.
- Avverka först de tre stegen snabbt och översiktligt. Bedöm därefter med hjälp av följande avbrottskriterier om stegen bör genomföras ytterligare en gång:
- Om det är osäkert om programmet uppfyller användarens behov – Repetera 1.2.3 I-III mer utförligt.
 - Om programmet kan uppfylla användarens behov – Gå vidare till 1.2.4.

- Om programmet inte tycks uppfylla användarens behov – Avsluta utredningen, eller undersök om de alternativa programvarorna (från 1.2.2) kan uppfylla behoven istället.

Om alternativet ”avsluta utredningen” väljs kan det vara lämpligt att försöka uppskatta de risker som är förknippade med att *felaktigt* välja att *inte* använda/införskaffa programvaran.

1.2.4 Steg 4 Fördjupad undersökning

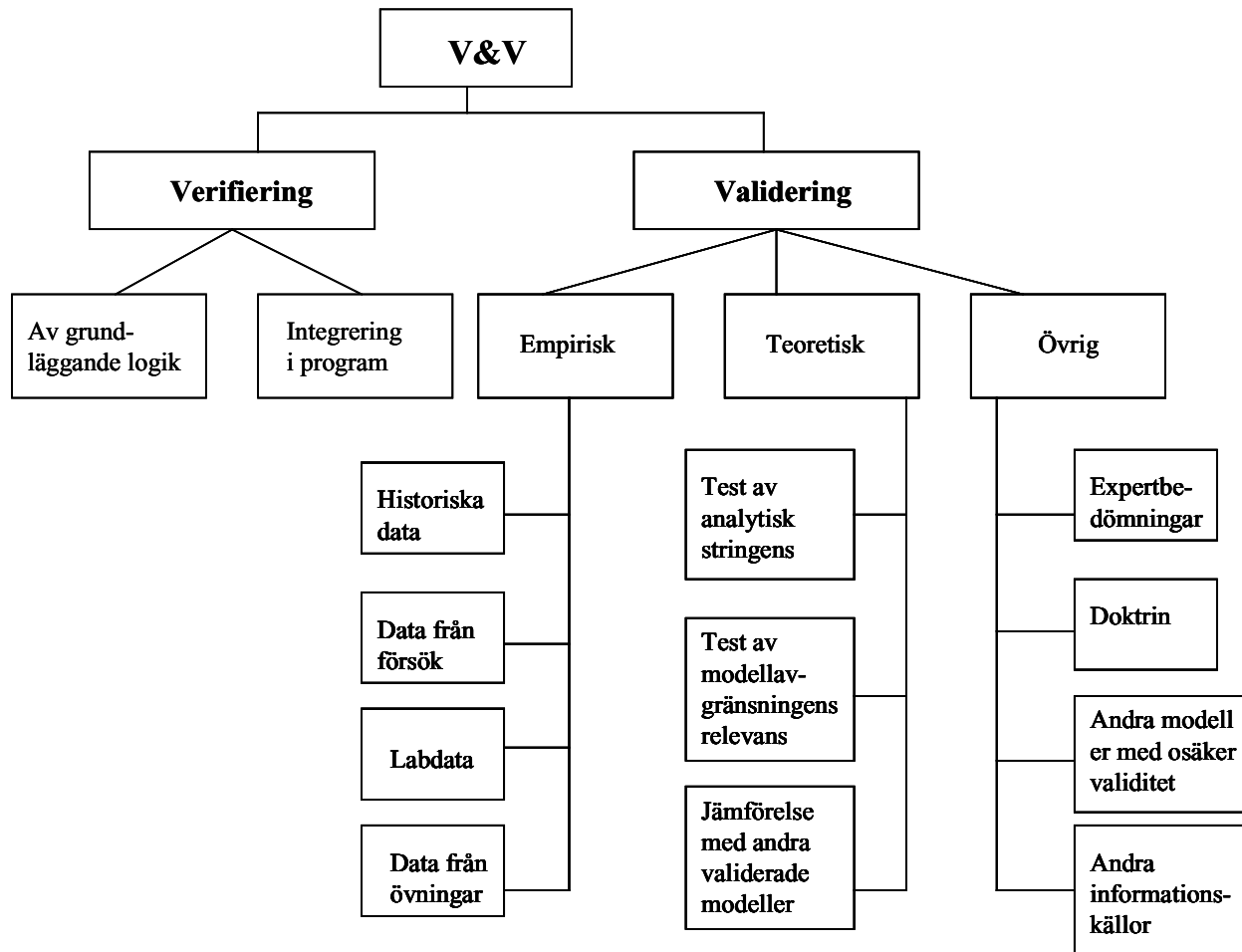
Om programmet bedöms kunna uppfylla en potentiell användares behov undersöks programmets trovärdighet, användbarhet och flexibilitet samt de kostnader som kan förknippas med bruket av programmet. Arbetet sker enligt följande:

- Validering och verifiering av resultat
- Undersökning av kvaliteten på presentationen av resultatet
- Undersökning av programmets genomskinlighet
- Undersökning av programmets flexibilitet
- Undersökning av tillgången på användarstöd
- Undersökning av kostnader

1.2.5 Steg 5 Dokumentation

Sammanfatta och presentera insamlat material.

Bilaga 2. V&V-metodik inom MoS



Figur 4 En översikt över V&V-metodiken. Förenklat från André *et al* (1999).

André *et al* (1999) har gjort en översikt över olika valideringsmetoder. Som framgår av figuren kan valideringen delas in i empirisk, teoretisk och övrig validering. Empirisk validering sker mot olika typer av data, t ex historiska data eller data från fältförsök och övningar. Teoretisk validering sker mot andra simuleringar och bedömningar av bland annat modellens analytiska stringens. Övriga valideringsmetoder utnyttjar experter och skrivet material. Dessa begrepp kan jämföras med validering mot testdata, "benchmarking" och "face validation" enligt Laack *et al* (2000). Bedömningar om modellens valideringsgrad görs utifrån resultaten från de olika metoderna.

Bilaga 3. Egenutveckling kontra inköp av modeller

Nedanstående sammanställning är bifogad för att visa hur en ”skarp” utvärdering av en befintlig modell kan se ut. Läsaren bör dock observera att utvärderingen inte skett enligt snabbsnigelmetoden men att de ändå har många gemensamma punkter.

Dokumentet sattes samman på FOA av Karin Mossberg, Tore Isacson, Eva Andersson, Susanne Odar, Tommy André och Erland Tarras-Wahlberg 1999. Bakgrunden till dokumentet var att FOA ville utveckla en egen luftkrigsmodell (med FLAMES) och FMV ville samarbeta med tyska IABG och ta del av deras modeller. HKV var uppdragsgivare och skulle ta beslutet. Listan gjordes för att tydliggöra för- och nackdelar med respektive alternativ.

Nedanstående punkter är inte listade i prioritetsordning utan har fyllts på efter hand som kommentarer inkommit.

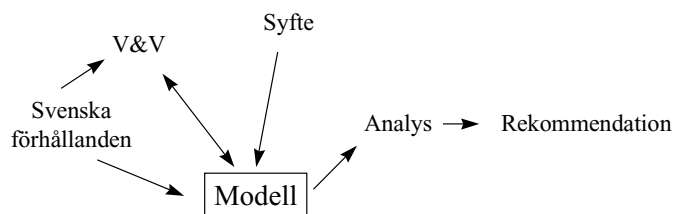
	Inköp av FLAMES, egenutveckling SMART	Inköp av tyska modeller, vidareutveckling
Kostnad utveckling	- Större*. Utvecklingen kostar mycket pengar och tar tid.	+ Mindre?*. Får vi modellen ”gratis” kostar inte utvecklingen oss någonting och vi får dessutom en färdig modell.
Underlagsframtagning	+ Indata till modellen måste tas fram. Modellen kan anpassas efter befintligt underlag och ambitionsnivån ökas efter hand.	- Indata till modellen måste tas fram. Indata styrs av befintlig modell och kan ej påverkas.
Förståelse för hur modellen fungerar och kan användas, analys av resultatet	+ Modellinsikt byggs upp under utvecklingen. Den måste dock bevaras vilket ställer krav på resurser.	- Modellinsikt måste köpas från IABG. Kan på lång sikt byggas upp på FOA men det kräver att resurser avsätts.
Anpassning av modellen efter syftet	+ Modellen kan ”skräddarsys” för det syfte den skall användas för.	- Vi får en modell, studierna får anpassas efter modellen.

Samarbete med IABG	+/- Samarbete kan (men måste inte) ske.	+/- Ett ökat samarbete är ett krav. Positiva bieffekter.
Vidmakthållning av modellen	+/-	+/-
Modifiering av modellen	+ Enklare att modifiera modellen vid egenutveckling.	- Kan vara svårt att modifiera modellen på egen hand. Kräver köp av tjänster från IABG. Det går ej att använda den telekrigssimulering som eventuellt finns i modellerna. Av sekretessskäl måste en svensk modell av TK införas. Det kan vara svårt att komplettera en befintlig modell.
Beroende av nyckelpersoner	+/- Större risk om MoS-verksamheten vid FOA bedrivs i liten skala.	+/- Nyckelpersonerna hamnar på IABG, även på FOA.
Löpande kostnader	+/- Underhållsavtal för FLAMES, 250.000 kr per år	+/- Resor mellan Tyskland och Sverige, ca 250.000 kr per år (?)
Simulering av svenska förhållanden	+ Byggs in under utvecklingen. Kan utvecklas för olika typer av scenarier.	- Måste modifieras i efterhand (kräver även förändringar i källkod, inte bara andra indata). Beroende av IABG. Vilken typ av scenarier stödjer modellen?
Dokumentation	+ FLAMES finns väl dokumenterat. Övrig utveckling dokumenterar vi själva.	-? Finns dokumentation för användare och utvecklare? Hur utförlig? Språk?
Validering	+/- Får vi göra själva. Detta ger en ökad insikt om vad modellen klarar av.	+/- Är de tyska modellerna validerade? Mot vad? Litar vi på valideringen?

Ackreditering	+ Kan göras utifrån den kunskap som byggs upp under utvecklandet.	- Ett stort arbete tillkommer för att ackreditera modellen efter vårt sätt att använda modellen.
Kompetens på FOA (FM) för att utveckla och använda modeller.	+ Byggs upp / vidmakthålls	- Ett beroende till IABG byggs upp.
Tid innan modellen kan användas	+/- Går att använda för att studera vissa problem redan idag. Fullt utvecklad om 1-3 år (beroende på finansiering)	+/- Går att demonstrera i närtid Använda i närtid? Det krävs dock att modellen valideras efter svenska mått och att indata tas fram. Kan "modellkunnig" person från IABG hjälpa till under skarpa Perp-körningar?
Beroende av utomstående (Ternion/IABG)	+ Ternion ger oss ett ramverk, scenariogenerering, grafik, databashantering, enkla modeller. Det som är modellspecifikt bygger vi själva).	- IABG ger oss hela modellen. Vi köper allt eller inget.
Modellstorlek	+ Kan anpassas till vad vi tror att vi kan hantera avseende indatabehov komplexitet vid modifieringar och insats för körningar	- Är ofta avpassad för att hanteras av en större organisation, d.v.s. hanterandet blir dyrare
Övertagande	+ Övertagande av modell är alltid ett problem, men underlättas om modellen inte är alltför stor, d.v.s. bra om den är anpassad för svenska förhållanden inte till någon annans + svenska förhållanden	- Troligen större modell, ett övertagande ingår redan i starten

Krigserfarenhet	- Finns begränsat inom landet	+ Kan finnas inbyggd i utländsk modell
Indata	+ Modellen kan anpassas till vår tillgång på indata	- Risk att kompetens för vissa indata måste byggas upp eller köpas
Krav på engagemang av beställaren (HKV alt. FlygvapenC)	+/- Stort om modellen ska anpassas efter syftet	+/- Mindre, men det innebär också ett mindre inflytande på modellens utformning.
Tidsperspektiv	+ Modellen kan utvecklas efter förhållanden år 2020.	- Modellen utformad efter system år 2000?
Programmeringsspråk	+ FLAMES är byggd på ny datateknik, stödjer ex. DIS och går att vidareutveckla i C++.	- Gamla program? FORTRAN? Stödjer ej DIS?
Storlek modeller	+ Går att utforma som många små modeller som kommunicerar med varandra. Är fördelaktigt då delar av modellen ska bytas ut.	- Stor monolit?

* Kostnaden för utveckling av en modell rymmer kostnader för analys, programutveckling, verifiering och validering samt delvis framtagning av underlag. För denna kostnad fås förutom en modell även en insikt om hur modellen fungerar och hur resultaten från den ska tolkas.



Verifiering och validering (V&V) sker parallellt med modellutvecklingen. Det innebär att en egenutvecklad modell är dynamisk och kan förändras beroende på utfallet av V&V. Att validera en färdigutvecklad statisk modell innebär alltid en risk att modellen inte uppfyller de krav som har ställts, och då finns ingen naturlig väg att förändra modellen.

