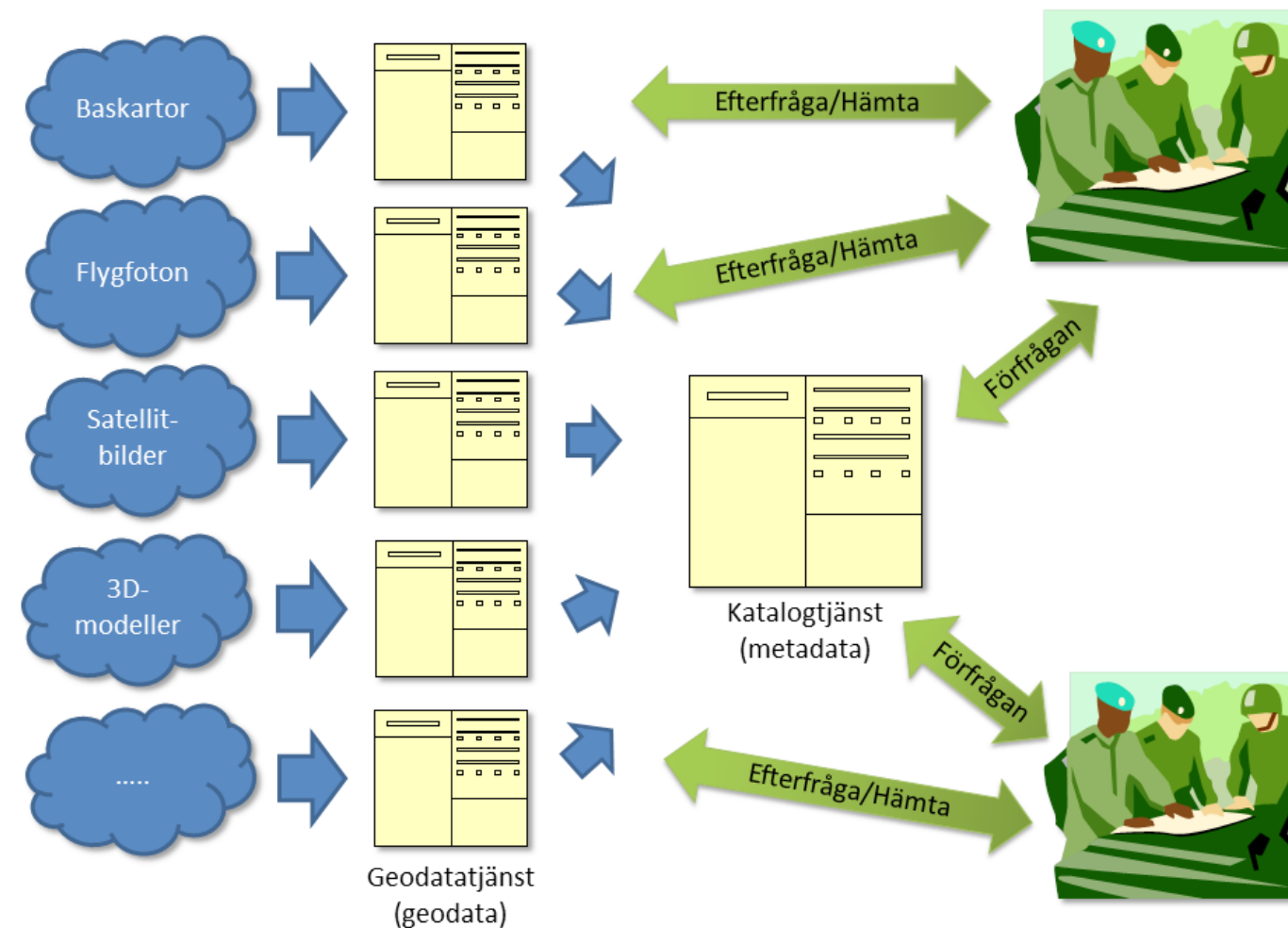


HANS-ÅKE OLSSON, PETER FOLLO,
JONAS HARALDSSON, HÅKAN LARSSON, GUSTAV TOLT



FOI är en huvudsakligen uppdragsfinansierad myndighet under Förvarsdepartementet. Kärnverksamheten är forskning, metod- och teknikutveckling till nytta för försvar och säkerhet. Organisationen har cirka 1000 anställda varav ungefär 800 är forskare. Detta gör organisationen till Sveriges största forskningsinstitut. FOI ger kunderna tillgång till ledande expertis inom ett stort antal tillämpningsområden såsom säkerhetspolitiska studier och analyser inom försvar och säkerhet, bedömning av olika typer av hot, system för ledning och hantering av kriser, skydd mot och hantering av farliga ämnen, IT-säkerhet och nya sensorers möjligheter.

Hans-Åke Olsson, Peter Follo, Jonas Haraldsson,
Håkan Larsson, Gustav Tolt

Metadata för 3D-informatik

Titel	Metadata för 3D-informatik
Title	Metadata for 3D Informatics
Rapportnr/Report no	FOI-R--3631--SE
Månad/Month	Januari
Utgivningsår/Year	2013
Antal sidor/Pages	56p
ISSN	1650-1942
Kund/Customer	FMV
FoT område	
Projektnr/Project no	E54068
Godkänd av/Approved by	Per Johannesson
Ansvarig avdelning	Sensor- och TK-system

Detta verk är skyddat enligt lagen (1960:729) om upphovsrätt till litterära och konstnärliga verk. All form av kopiering, översättning eller bearbetning utan medgivande är förbjuden.

This work is protected under the Act on Copyright in Literary and Artistic Works (SFS 1960:729). Any form of reproduction, translation or modification without permission is prohibited.

Sammanfattning

Metadata – ”data om data” – har till syfte att beskriva en dataresurs och är en viktig komponent för att geografiska data ska kunna få ökad spridning och acceptans. Metadata utgör grunden för katalogtjänster och hjälper användare att hitta data som passar i en viss tillämpning och att värdera data gentemot varandra. Nyttan med metadata kommer sannolikt att accentueras i takt med att utbudet av geografiska data ökar.

Existerande metadatastandarder för geografiska data är uttrycksfulla men samtidigt generellt skrivna. De saknar därför t ex explicit formaliserat stöd för att beskriva utmärkande egenskaper hos 3D/höjddata, såsom att noggrannhet kan vara olika i plan respektive höjd eller att det finns flera olika typer av höjdangivelser (t ex markens höjd, den högsta höjden eller djup i en sjö). Exempel på hur sådan information kan läggas i metadata är att använda beskrivande nyckelord eller att utforma en egen profil av standarden genom att t ex utöka den med nya eller ändrade listor över beskrivande kodord. Även om det senare alternativet begränsar uttrycksmöjligheterna, eliminerar det nackdelen med det godtycke som nyckelorden medför. Det är inte minst viktigt eftersom en viss begreppsförvirring råder på höjddataområdet.

Relevant kvalitetsrelaterad information i metadata är viktigt för att användare av data ska kunna bedöma hur pålitliga data är och avgöra till vad de kan användas. Detta är särskilt viktigt för nya användare med begränsade bakgrundskunskaper och ingen eller ringa erfarenhet av liknande data sedan tidigare.

Kvalitetsuppgifter kan även nyttjas i (GIS-)analyser, t ex för att kombinera flera heterogena dataset till förädlade produkter. Genom att ta hänsyn till respektive datalagers kvalitet (t ex noggrannhet i plan) kan en förfinad analys göras och resultaten värderas på ett annat sätt än om kvalitetsinformation saknas. Kända osäkerheter i data kan spåras genom analysprocessen och ge användare och beslutsfattare bättre underlag. I den här rapporten exemplifieras detta genom ett scenario som går ut på att detektera möjliga helikopterlandningsplatser utifrån olika informationskällor (höjddata och flygbilder). Exemplet illustrerar det mer generella problemet med hantering av heterogena dataset.

För att få svar på frågor kring vilka uppgifter som bör återfinnas i metadata, kan en strukturerad behovs- och kravanalys göras. I rapporten visas hur man utifrån systemutvecklingsmetoder kan få fram dessa uppgifter, exemplifieras med det ovan nämnda helikopterlandningsfallet.

Nyckelord: Metadata för 3D-produkter, höjddata, sökbarhet, kvalitet, INSPIRE, ISO 19115:2003, ISO 19139:2007, ISO 19115-2, ISO 19115-1, referensdata

Summary

Metadata – “data about data” – aim at describing a data resource and it is an important factor for increasing the exploitation and acceptance of geographical data. It forms a basis for catalogue services and helps users find data suitable for a particular application and assess different data sets. The benefits of metadata will be accentuated as the amount and variety of geographical data increase.

Existing metadata standards for geographic data are expressive, yet formulated in a quite general manner. As such, they lack explicit formalized support for describing typical characteristics of 3D/elevation data, e.g., that vertical and horizontal uncertainty can differ or that there is a range of different elevation data types (e.g., the elevation of the bare earth, the highest elevation, bathymetric depth, etc). Such information can be added to the metadata through descriptive keywords or through extending the standard with new or modified code lists of allowed code words. Although using pre-defined code lists limits the expressive power, it eliminates the disadvantages associated with the arbitrariness of keywords. This is an important point, as some confusion seems to exist in this field concerning the interpretation of elevation-related terms.

Relevant quality-related information in metadata helps users judge the reliability of data and determine for what the data could be used. This is particularly important for new users with limited knowledge and no or little previous experience of similar data.

Quality information can also be exploited in (GIS) data analyses, e.g., in order to combine multiple heterogeneous data sets into refined products. By taking the quality of the respective data layers into account (e.g., horizontal accuracy), a more detailed analysis can be performed, improving the possibility to interpret the result of the analysis. Known uncertainties in data can be tracked through the analysis process, in order to give users and commanders a more solid basis for decision-making. In the report, this is demonstrated in a scenario where possible helicopter landing zones are automatically detected by combining different sources of information (elevation data and aerial photographs). This example illustrates the more general problem of handling heterogeneous data sets.

In order to determine what information metadata should contain, a structured analysis of user needs and associated requirements can be performed. In the report, shows how these specifications could be produced by system development methods, exemplified by the helicopter landing zones case outlined above.

Keywords: Metadata for 3D products, elevation, searchability, quality, INSPIRE, ISO 19115:2003, ISO 19139:2007, ISO 19115-2, ISO 19115-1, reference data

Innehållsförteckning

1	Inledning	7
1.1	Om projektet.....	7
1.2	Rapportens upplägg.....	8
1.3	Flödet för geo- och metadata från behov till tillämpning	9
1.3.1	Inventering av behov och krav för att identifiera geo-, meta- och övriga data för en tillämpning	10
1.3.2	Framtagning av geo-, meta- och övriga data.....	10
1.3.3	Leverans av lämplig geo- och metadata för en tillämpning	11
2	Metadata på geoområdet	12
2.1	Begreppet metadata	12
2.2	Nyttan med metadata.....	12
2.2.1	Bredare användning av geodata.....	13
2.2.2	Tillförlitlighet	13
2.2.3	Hitta rätt data via söktjänster	13
2.2.4	Maskinell hantering	13
2.3	Standardiseringsarbete i världen	14
2.3.1	ISO	14
2.3.2	OGC	15
2.3.3	EU/INSPIRE.....	16
2.3.4	Sverige	16
2.3.5	DGIWG.....	17
2.3.6	NATO	17
2.3.7	USA.....	17
3	Metadata för 3D-produkter	19
3.1	Hur ser det ut idag?	19
3.2	Vilka metadata behövs för 3D-produkter?	19
3.2.1	Översiktlig beskrivning	19
3.2.2	Referenssystem	20
3.2.3	Typ av höjdangivelse	20
3.2.4	Ursprung och processhistorik	21
3.2.5	Täckning	21
3.2.6	Kvalitetsmått	21
3.2.7	Kvalitetsvärderingsmetod.....	23

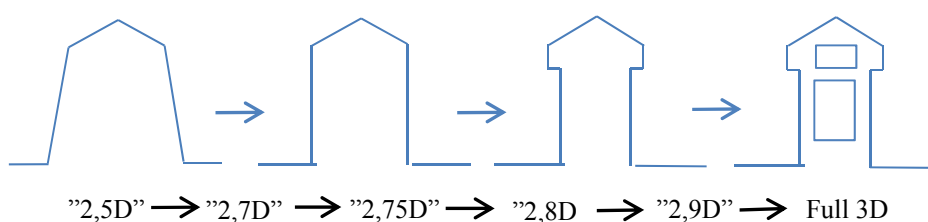
3.2.8	Upplösning.....	23
3.2.9	Aggregerade dataset.....	24
4	Hur kan man skapa bättre metadata	25
4.1	Behovsinventering	25
4.2	Krav på leverantörer	25
4.3	Standarder	25
4.4	Process för kvalitetsvärdering av 3D/höjddata	26
5	Referensdata	27
5.1	Data	27
5.1.1	NNH	27
5.1.2	Högupplöst flygburen laserskanning	28
5.1.3	R3DM.....	28
5.1.4	Markbunden 3D Laserskanner	29
5.2	Metadata för referensdata	31
6	Testfall/scenario	32
6.1	Beskrivning av scenariot.....	32
6.1.1	Reglemente	33
6.1.2	Produktspecifikation	34
6.1.3	Metod.....	34
6.1.4	Ingående datalager.....	35
6.1.5	Felfortplantning.....	38
7	Kontakter	42
8	Förslag på fortsatt arbete	43
9	Referenser	44
10	Förkortningar	45
Bilaga 1		47
	Identifiera och_kravställa geoinformation med tillhörande metadata.....	47

1 Inledning

1.1 Om projektet

Projektet är en beställning från FMV inom ramen för FMV:s projekt SGIF (Snabb Geografisk InformationsFörsörjning) och syftar till att kunna beskriva "3D-data" genom metadata så att dessa data kan användas säkert i olika Försvarsmaktstillämpningar. Projektet har särskilt beaktat hur metadata avseende kvalitetsaspekter kan användas för att stödja förståelse av hur "pålitligt" data är och därmed ge ökad säkerhet och trygghet vid användning av 3D-geoprodukter inom geoinfotjänsten.

Målet för projektet har varit att (1) genomföra en teoretisk och praktisk studie om metadataproblematiken för 3D-produkter och (2) en demonstration av ett första koncept för metadatahantering för 3D-produkter samt (3) att etablera 3D referensdata för utvalda delområden. En fråga att adressera tidigt är vad som egentligen avses med "3D-data" och "3D-produkter". Betydelsen av begreppet "3D" är ganska luddigt och kan variera från person till person. Den gemensamma nämnaren för "3D"-data, som skiljer dem från "2D"-data, är dock att de bär information om terrängens variation i alla tre dimensionerna (x,y,z) och inte bara i planet (x,y). Inom den stora gruppen finns sedan en rad olika varianter. Verklig (äkte, full) 3D har man endast om representationen tillåter att godtyckliga geometriska strukturer kan beskrivas. Terrängens höjdvariationer kan emellertid beskrivas på sätt som inte nödvändigtvis inbegriper en tredimensionell funktion. Ur en matematisk synvinkel är t ex en vanlig höjdbild inget annat än en tvådimensionell representation, eftersom höjden (typiskt "z") varierar som funktion av koordinaterna (x,y). Detsamma gäller en höjdrepresentation i form av höjdkurvor. För att förtydliga att funktionsvärdet representerar höjd, brukar höjdbilder ofta kallas "2,5D". En nackdel med höjdbilder är att den endast har ett höjdvärde och därmed inte kan representera vertikala strukturer eller överhäng (t ex broar). Rapportförfattarna har i olika sammanhang stött på de minst sagt luddiga begreppen "2,7D", "2,75D", "2,8D" och "2,9D" som används för att beskriva representation som bättre kan beskriva den verkliga geometrin men som ändå inte förtjänar att kallas "3D". Det kan exempelvis handla om representationer som tillåter vertikala strukturer eller olika grader av överhäng, se Figur 1. Som begrepp är dock dessa inte lika vanligt förekommande som "2,5D".



Figur 1. Illustration av olika förekommande begrepp kopplade till höjd/3D-data. (Läsaren ombeds tänka sig figurerna som sedda från sidan, dvs som tvärsnitt av modeller som sträcker sig inåt i papperet).

I det här projektet tolkar vi begreppet "3D" väldigt brett och låter det omfatta alla dessa olika varianter på höjddata, eftersom vårt intresse inte ligger i själva representationen av data utan hur data skall beskrivas genom metadata.

Vägledande frågor för projektet har bl.a. varit:

- Vilka metadata behövs för 3D-produkter och deras tillämpningar?
- Hur kan metadata representeras och kodas på ett maskinläsbart sätt för att stödja användning av 3D-produkter i automatiska analyser?

- Kan varierande datakvalitet visualiseras så att en beslutsfattare vet var data är mer pålitligt respektive mindre pålitligt?
- Vad händer med metadata när nya 3D-geodata skapas genom bearbetning av andra 3D-geodata och hur kan tillhörande metadata skapas?
- Vilka höjd/3D-data finns idag tillgängliga för experimentella studier för ovanstående?
- Behöver existerande data kompletteras med extra noggranna referensmätningar, t ex mha markbunden laserskanning?

Arbetet har genomförts i fem arbetspaket:

- AP1.1 – Beskriva Metadataproblematiken
 - Beskriva vikten av och problematiken kring geometadata
 - Litteraturstudie
 - Beskriva arbete med standarder (nationellt/internationellt)
- AP1.2 – Metadata för 3D-data
 - Undersöka standarder och praxis avseende metadata inom 3D-området
 - Ta fram metadatakoncept för en eller flera 3D-produkter, i dialog med beställaren och/eller slutanvändare
- AP1.3 – Referensdata
 - Etablera referensdataset för experiment och demonstration
 - Genomföra kompletterande inmätningar efter behov
- AP1.4 – Experiment
 - Utredda tekniska förutsättningar för maskinell hantering och visualisering av metadata
 - Teknisk implementation
 - Värdering av koncept tillsammans med beställaren och/eller slutanvändare
- AP1.5 – Demonstration
 - Planera och genomföra demonstration

Projektet har, förutom denna rapport, avrapporterats i två statusrapporter. Dock har demonstrationen, efter överenskommelse med FMV, flyttats fram i tiden och avrapporteras därför inte här.

1.2 Rapportens upplägg

Rapporten inleds med en kort översikt över hur projektet har varit inriktat, exempel på hur flödet för geo- och metadata från behov till tillämpning kan se ut för att ge en övergripande inledning.

Därefter beskrivs begreppet metadata och den nytta man kan ha av bra metadata inom geoområdet igenom och en översikt ges av delar av det standardiseringsarbete som sker nationellt och internationellt inom geoområdet. Metadata specifikt för 3D-produkter diskuteras igenom; hur ser det ut idag och vilka metadata behövs för 3D-produkter.

Då metadata idag inte alltid är så omfattande och så kvalitativt bra som vore önskvärt i olika tillämpningar, blir frågan hur man då skapa bättre metadata. I rapporten ges exempel på vad som kan göras för att få rätt kan metadata med rätt kvalitet genom t ex behovsinventering, kravställning på leverantörer och nyttjande av standarder.

Projektet har byggt upp ett höjd/3D-data referensdataset för det experimentella arbetet i projektet. Datat är över Strid-i-Bebyggelse-anläggningen på Markstridsskolan i Kvarn. Rapporten innehåller en redogörelse av de ingående komponenterna i detta referensdataset.

För att åskådliggöra nyttan med metadata som är kopplad till geoinformation och hur metadata kan användas, har ett fiktivt scenario för en helikopterlandningsplats tagits fram. Med hjälp av metadata, analyseras geodata (inkluderande 3D data) och övrig data för att få fram lämpliga landningsplatser samt för att bedöma möjliga felfortplantningar som kan påverka resultatet. Detta beskrivs och de olika stegen och resultaten åskådliggörs.

Rapporten avslutas med förslag på fortsatt arbetet utifrån projektets uppbyggda kunskap och erfarenheter. En bilaga bifogas, som visar hur man kan arbeta utifrån systemutvecklingsmetoder för att åstadkomma en spårbar koppling mellan slutanvändarnas behov till de faktiska kraven på geodata med tillhörande metadata.

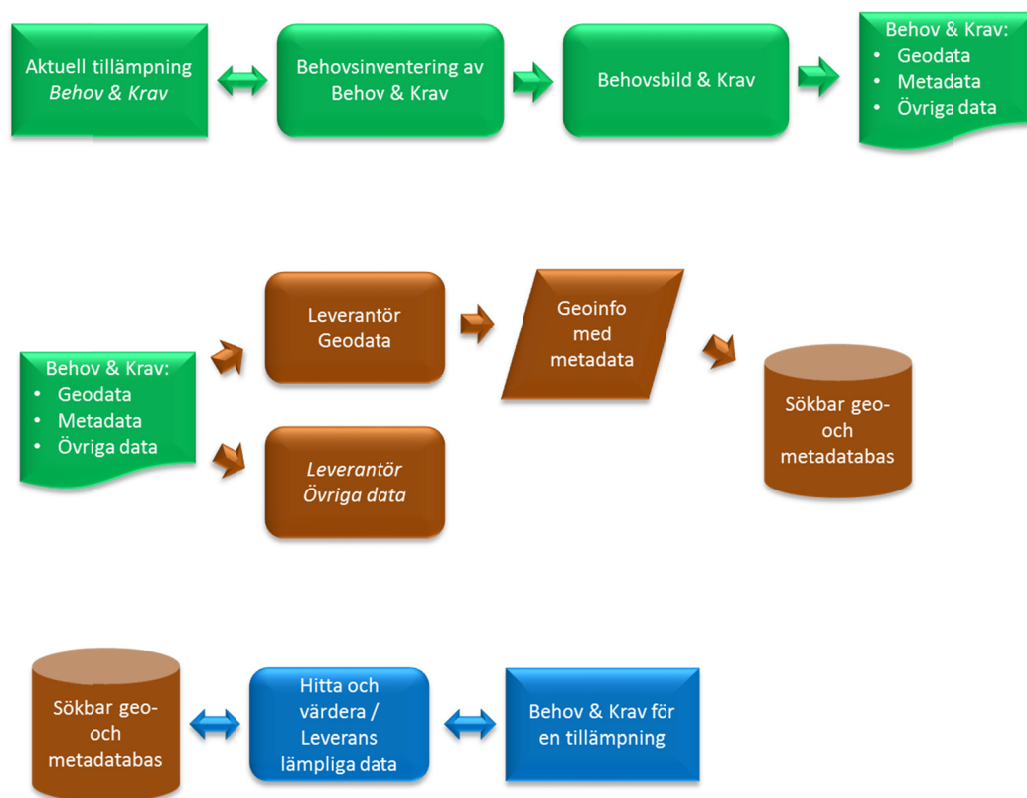
1.3 Flödet för geo- och metadata från behov till tillämpning

Denna rapport beskriver problematiken kring metadata för geoinformation med särskilt fokus på 3D-data.

Metadata är en del i ett geodataset och beskriver den geodatamängd som utgör den andra delen i ett geodataset. Exempel på geodata kan vara satellitbilder, flygbilder och laserhöjdmätningar. Ett geodataset ingår i en produkt eller tjänst för att t ex hitta lämpliga helikopterlandningsplatser, göra siktanalyser och beräkna riskområden.

Man kan beskriva flödet för geo- och metadata från behov till tillämpning i tre steg (Figur 2):

1. Inventering av behov och krav för att identifiera geo-, meta- och övriga data för en tillämpning
2. Framtagning av geo-, meta- och övriga data
3. Leverans av lämplig geo- och metadata för en tillämpning



Figur 2. Flödet för geo- och metadata från behov till tillämpning

1.3.1 Inventering av behov och krav för att identifiera geo-, meta- och övriga data för en tillämpning

Vilka metadata och kvalitén på dessa som behövs till en datamängd beror till stor del på de krav som tillämpningen ställer på geodata där dessa skall användas. D v s det behövs kännedom om vilka tillämpningar/verksamheter som datamängden kommer att användas i, så att en spårbar koppling mellan slutanvändarnas behov till de faktiska kraven på metadata kan erhållas.

Förutom rätt geo- och metadata kan tillämpningen också ställa krav på andra data som t ex ROE och tekniska specifikationer. I denna rapport avhandlas bara geo- och metadata.

Genom att göra en behovsinventering, kan man erhålla vilka behov och krav som en tillämpning ställer. Utifrån detta går det sen att få fram vilka geodata och övriga data som behövs för tillämpningen, samt hur metadata för dessa bör se ut. Denna inventering ger också en ökad förståelse för när olika geodata kan användas. Se vidare kap. 4.

1.3.2 Framtagning av geo-, meta- och övriga data

När väl behovet av geo-, meta- och övriga data har fastställts för en tillämpning, kan anskaffningen av data påbörjas. Det framkomna behovet av geo- och metadata kan fungera som en kravspecifikation när upphandling av geodata görs. I kraven bör också metadatakrav från andra relevanta tillämpningar tas med. Detta för att få så många metadata parametrar som möjligt till datasetet och därmed få ett så stort användningsområde som möjligt för de levererade geodata.

Det framkomna behovet av metadata och kvalitetsmått för dessa, utgör en bra kravställning gentemot leverantören. Dels att data bör uppfylla dessa krav och dels att

leverantören måste ange vilka krav man uppfyller och specificera hur man har säkerställt att kraven uppfylls.

Det bör också ställas krav på att de levererade metadata filerna följer lämplig ISO standard, detta för att möjliggöra automatisk inläsning av metadata i olika tillämpningar.

När data sen levereras, så har man säkerställt att man har fått fram så mycket information som det är möjligt.

Genom att systematisk arbeta igenom behov och krav för tillämpningar och utifrån detta ställa krav på leverantören, har man fått en bättre beskrivning och bättre kvalité av data. Detta ger fler användningsområden där data kan användas tryggt inom angivna säkerhetskrav.

Levererade data kan sen lagras i en databas, där geo- och metadata är sökbara för andra tillämpningar samt för att skapa olika produkter och tjänster.

1.3.3 Leverans av lämplig geo- och metadata för en tillämpning

Ju mer metadata som finns till geodata, desto större möjlighet finns det att dessa data också kan användas säkert för andra tillämpningar än vad som var tänkt ifrån början. Genom en bra beskrivning av geo- och metadata blir det lättare att hitta data och att värdera olika dataset maskinellt. Då både för att identifiera om det finns lämpliga dataset som skulle kunna används och för att kunna värdera olika dataset gentemot varandra.

2 Metadata på geoområdet

En av nycklarna till ökad och förbättrad användning av geografiska data i olika tillämpningar ligger i att etablera en genomtänkt och effektiv process för hantering av metadata. Metadata beskrivs ofta som ”data om data” och syftar till att ge datamängder på olika nivåer meningsfulla beskrivningar som ger förutsättningar för att data kan hanteras på ”rätt” sätt.

Man brukar ofta skilja på innehåll och representation. Det innehållsmässiga handlar t ex om vilka egenskaper hos dataresursen som ska beskrivas, hur de hör ihop rent logiskt, vilken information som är obligatorisk och vilken som är frivillig, etc. Metadata för geografiska data kan t ex innehålla uppgifter om vem som producerat data, när och hur data har samlats in, i vilket syfte, datas kvalitet, geografiskt område, koordinatsystem, dataformat, etc. Den andra delen handlar om hur innehållet ska kodas, t ex vilket filformat som ska användas.

2.1 Begreppet metadata

Generellt sett är metadata applicerbart på alla hierarkiska nivåer, från de lägsta (enskilda mätpunkter eller pixlar) till de högsta (produktsviter eller hela kartserier). Det är därför ofta tämligen meningslöst att diskutera vad som är metadata och vad som inte är det. Den som definierar en geodataresurs – oavsett datanivå – sätter (indirekt) ramarna för vad metadata bör beskriva. Hur långt ner i hierarkin det är lämpligt att gå avgörs typiskt av faktorer som t ex datalagring, ajourhållning, hur mycket information som finns och hur den skiljer sig från andra jämförbara objekt, etc. Även om det är teoretiskt möjligt att ha en utförlig metadatafil för varje pixel i en bild är det knappast relevant i särskilt många tillämpningar eller ens praktiskt möjligt att hantera.

Ju längre ner man kommer i datahierarkin, desto vanligare blir det att man stöter på andra begrepp än just metadata. På lägre nivåer (t ex för enskilda objekt) brukar man istället för metadata ofta tala om attribut eller egenskaper för dataobjekten, även fast de också kan ses som ett slags metadata. Inom INSPIRE (se avsnitt 2.3.3) gör man just denna observation¹. Den högsta metadata nivån kallas där *dataset-level* medan den objektsspecifika benämns *spatial object-level*.

Metadata kan alltså konstrueras på olika aggregationsnivåer och bör hela tiden beskriva dataresursen på den aktuella nivån. Om en dataresurs består av ett sammansatt kartpaket bör det finnas metadata för var och en av de ingående komponenterna och över denna nivå metadata som beskriver den sammansatta dataresursen.

Begreppet metadata kan också delas upp i olika klasser och underklasser men vi avser här inte gå djupt in på nomenklaturen inom området. Vi ämnar således inte utreda skillnader mellan olika i forskningslitteraturen föreslagna metadata typer som t ex strukturella, tekniska, deskriptiva och administrativa metadata utan antar ett mer pragmatiskt förhållningssätt fokuserat på Försvarmaktens behov och verksamhet inom geoområdet.

2.2 Nyttan med metadata

Vad har man då för nytta av (bra) metadata?

¹ “D2.8.II.1 Data Specification on Elevation – Draft Guidelines”, INSPIRE Thematic Working Group Elevation, 2011-06-15

2.2.1 Bredare användning av geodata

En bra metadatabeskrivning ger bättre förutsättningar för att använda geodata i olika tillämpningar. Genom att specificera centrala egenskaper hos data och beskriva dessa i metadata förbättras möjligheterna till att hitta rätt data för en viss tillämpning, men också för det omvända: att hitta nya tillämpningar för ett visst dataset som t ex går utanför vad data ursprungligen tagits fram för. Kopplingen mellan behov och krav på data blir med andra ord tydligare om man väljer data utifrån t ex kvalitetsrelaterade mått (noggrannhet, precision, etc) istället för att låsa sig vid data från en viss leverantör eller på ett visst format.

En situation där metadata saknas eller är väldigt bristfällig kräver i praktiken att användare av data har stor kunskap och erfarenhet av liknande data. Bra metadata är därför i princip helt nödvändig för att geodatabaserade analyser och bedömningar ska kunna utföras av personer utan större GIS-bakgrund och därigenom få större spridning. Metadata är naturligtvis inte ett fullgott substitut för fackkunskap och erfarenheter, men tjänar likväl till sprida och öka förståelsen för geodata i olika tillämpningar.

2.2.2 Tillförlitlighet

Ökad och mer avancerad användning av geodata väcker frågor kring tillförlitlighet. En rik metadatabeskrivning ger beslutsfattare kunskap om vilket underlag han eller hon har att agera efter. Det kan t ex handla om faktorer som olika kvalitetsmått, uppgifter om processhistorik, t ex vilka förädlingssteg som det aktuella datasetet genomgått eller vilka dataset som används för att framställa datasetet. Information om hur kvaliteten fastställts kan också vara väldigt relevant. Har data genomgått en rigorös kvalitetssäkring enligt någon välkänd metod, eller har någon "gissat" vad data håller för kvalitet?

Detta kopplar också till *semantisk interoperabilitet*, som i korthet betyder att olika användare/system har samma förståelse för olika begrepps betydelse. Där har metadata en viktig roll att spela, eftersom det är ett redskap för att specificera vad data är och därigenom minska risken för missförstånd mellan användare. Detta kopplar i sin tur till frågor om standardiserad nomenklatur för metadata.

2.2.3 Hitta rätt data via söktjänster

Metadata möjliggör också publicering av data så att andra användare kan hitta och använda dessa. Det kanske mest intressanta exemplet på det är standarden CSW (*Catalog Service for the Web*) som tagits fram inom Open Geospatial Consortium (OGC). Standarden definierar enhetliga gränssnitt för metadata för geodata, men även för tjänster (WMS, WFS, WCS, etc). De flesta större GIS-leverantörer har numera stöd för detta och relevansen för denna typ av tjänster blir större i takt med geoutbudet på Internet ökar. Man skulle kunna säga att CSW har samma roll för geodata(tjänster) som en sökmotor på Internet har för webbsidor och deras innehåll.

2.2.4 Maskinell hantering

Metadata är också en nyckel till automatisk, maskinell hantering av geodata. I metadata finns (ska finnas) information som gör det möjligt att läsa in, analysera och manipulera data i en automatisk process. Uppgifter i metadata kan t ex användas i processen för att relatera dataresursen till rätt geografisk position mha information om koordinatsystem, göra felfortplantningsanalyser utifrån olika kvalitetsmått, avgöra dataresursens lämplighet för en viss analys utifrån dess klassificering/typ, etc.

En fråga som aktualiserats på senare tid rör hantering av *heterogena* dataset. Det handlar om att kunna kombinera flera underlag som skiljer sig åt med avseende på upplösning, koordinatsystem, geografisk utbredning, insamlingsmetod, eller insamlingstidpunkt, etc. Här kan metadata användas för att skapa bästa tänkbara förutsättningar för att en

användare ska kunna utnyttja data på rätt sätt i en viss tillämpning. Det kan t ex handla om samtidig visualisering av flera dataresurser för avdömning av deras respektive förtjänster och brister, fusionera data genom att låta den i något avseende bästa dataresursen användas för beräkningar för ett visst område, att göra bedömningar som baseras på olika typer av dataresurser, etc.

2.3 Standardiseringsarbete i världen

På geodataområdet finns idag en lång rad olika initiativ, nationella och internationella, som reglerar såväl innehåll i som teknisk implementation av metadata inom olika organisationer. Initiativen grundar sig ofta på någon eller några existerande standarder men måste ofta anpassas för att tillmötesgå specifika egna behov inom den egna organisationen eller behov av samarbeten mellan aktörer. Nedan följer en kort sammanställning av några viktiga initiativ inom området.

2.3.1 ISO

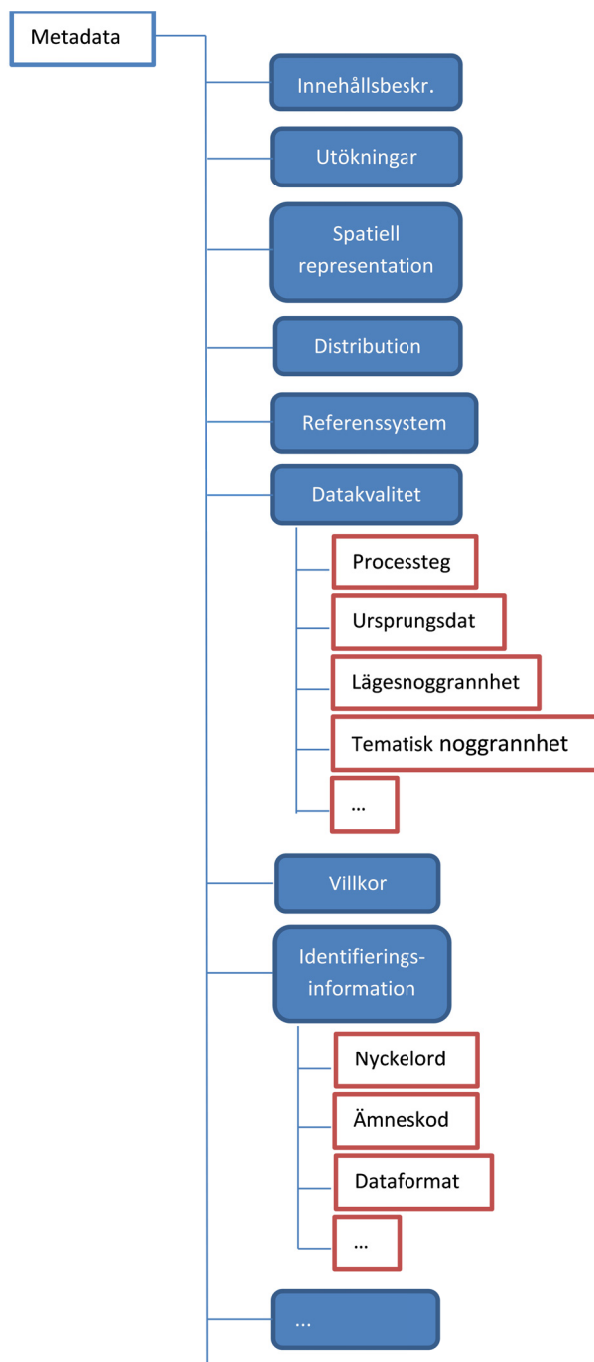
The International Organization for Standardization (ISO) har tagit fram flera standarder för geografisk information (ISOs 19100-serie). ISO 19111 handlar t ex om rumsliga referenser mha koordinater, ISO 19113 definierar kvalitetsprinciper för geografiska data, ISO 19114 specificerar utvärderingsprocedurer för dessa, osv. För själva metadata är det särskilt två standarder som tjänar som fundament i mycket av arbetet inom andra organisationer och forum:

- *ISO 19115:2003* (Geospatial information - Metadata) som reglerar vilken information metadata ska innehålla och hur den ska organiseras,
- *ISO 19139:2007* (Geospatial information – Metadata – XML schema implementation) som specificerar hur metadatainformationen ska kodas i form av XML för att få metadata på maskinläsbar form.

Dessa standarder bygger på flera andra standarder i 19100-serien, så närmare förklaringar av termer i dessa står ofta att finna i andra dokument.

Det finns sedan 2010 en utökning till ISO 19115 för bilder och rasterade ("griddade") data som benämns ISO 19115-2. Utökningen syftar till att tillföra ytterligare bildrelaterade detaljer jämfört med 19115, bl a angående sensorn, datainsamlingsprocessen och databearbetningen liksom information om eventuella kontrollpunkter som använts för kvalitetsbedömning. Utökningen ger dessutom större möjligheter att uttrycka datas kvalitet i termer av lämplighet i en viss tillämpning, som ett komplement till de i ISO 19115 specificerade kvalitetsmåten. För exempel på element som ingår i ISO 19115 (Figur 3).

I skrivande stund pågår dessutom arbete med att ta fram en reviderad version av ISO 19115, benämnd ISO 19115-1. Förutom att flera element kommer att modifieras och tillkomma väntas den nya versionen inkludera tjänster för metadata (från ISO 19119). Den nya internationella standarden väntas bli accepterad under 2013. I kölvattnet av denna kommer en ny XML-implementation tas fram (på samma sätt som ISO 19139 implementerar ISO 19115).



Figur 3. Ett urval av element som ingår i standarden ISO 19115. Standarden innehåller några hundra element indelade i olika grupper. En del av elementen är obligatoriska, medan andra är rekommenderade eller obligatoriska under vissa betingelser.

2.3.2 OGC

The Open Geospatial Consortium, OGC, är ett internationellt konsortium bestående av för närvarande 482 företag, statliga organisationer och universitet som tar fram georelaterade standarder. OGCs syfte är att förenkla distribution och hantering av olika typer av rumslig information och därtill hörande tjänster. OGC arbetar enligt en konsensusprocess för att få så stor spridning och interoperabilitet som möjligt för sina standarder. OGCs portfölj inkluderar bl a *Geography Markup Language* (GML), CityGML, webbtjänsterna WMS, WFS, WCS, WPS, m fl, liksom tjänster för hantering av sensordata. För metadata finns

katalogtjänsten CSW som ett verktyg att publicera och komma åt metadata på ett standardiserat sätt över nätet.

2.3.3 EU/INSPIRE

EU-direktivet INSPIRE² (*Infrastructure for Spatial Information in the European Community*) syftar till att få till stånd en etablering av en infrastruktur för rumslig information för att stödja miljörelaterat arbete inom EU. Eftersom stora delar av den totala mängden geografisk information har någon sorts bäring på miljöfrågor, har detta direktiv lett till en snabb teknisk utveckling på geområdet i stort i Europa. Exempelvis framtagande av en mängd metadatauppgifter med vars hjälp data ska beskrivas och utveckling av portaler där man genom sökningar i metadata kan hitta dataresurser och ibland även ladda ner dem.

Inom INSPIRE pågår dessutom just nu arbete med att ta fram en dataspecifikation för höjddata som behandlar både data och metadata. I det senaste dokumentet³ rekommenderas bl a att förstahandsvalet för höjd- (liksom djup-) data ska vara GML, men att höjddata i rasterform även kan få representeras på andra inom området vanligt förekommande format förutsatt att ingen förstörande datakompression förekommer. För höjddata som representeras som ett TIN (*Triangulated Irregular Network*) föreslås att OGC-standarden CityGML användas. Dessutom föreslås en rad metadatauppgifter relaterade till datas kvalitet, geografiska utsträckning, etc. INSPIREs metadatamodell bygger på ISO 19115 och ISO 19119.

Det återstår att se vilken påverkan den slutliga specifikationen kommer att ha på geodatahantering kopplade till Försvarmaktens behov. Försvarmkten har ett intresse av att tillgodogöra sig utvecklingen på det europeiska planet men strävar samtidigt efter att vara interoperabla med partners i samband med operationer runt om i världen. Man kan i och för sig omforma och konvertera data mellan olika format, men till priset av att all metadatainformation i INSPIREs dataspecifikation kanske inte går att nyttiggöra sig i ett system som använder data på ISO-struktur.

2.3.4 Sverige

Varje medlemsland i EU är skyldigt att vidta åtgärder för att tillgodose behovet av åtkomst till miljörelaterade geodata enligt INSPIRE-direktivet. I Sverige har detta bl a lett till införandet av Geodataportalen⁴. Data som publiceras på portalen måste beskrivas med metadata enligt en Nationell metadataprofil⁵ (SIS/TK 489 N247). Denna profil är en tillämpning av standarderna ISO 19115 och ISO 19139 och följer i stort INSPIRE-profilen, med några ändringar⁶ som ansetts viktiga för nationell geodatasamverkan.

Lantmäteriet har fått regeringens uppdrag att framställa en Ny Nationell Höjdmodell (NNH). Framställningen pågår och ca 2/3 av landet har redan täckts in. Till levererade data finns knutet en XML-metadatafil⁷. Denna fil tycks inte följa någon standardiserad metadataprofil utan är mer strömlinjeformad för att tillgodose behov kopplade till NNH, t ex med information om noggrannhet i höjddata, uppgifter om programvaruversion för datagriddning och databearbetning, etc.

² <http://inspire.jrc.ec.europa.eu/>

³ "D2.8.II.1 Data Specification on Elevation – Draft Guidelines", INSPIRE Thematic Working Group Elevation, 2012-07-09

⁴ <http://www.geodata.se/>

⁵ <http://www.geodata.se/Vad/Om-Geodataportalen/Nationell-metadataprofil/>

⁶ Några metadataelement som är frivilliga att tillhandahålla i INSPIRE-direktivet är obligatoriska i den nationella profilen

⁷ <http://namespace.lantmateriet.se/gdshojd/>

På Försvarsmaktens stödenhet för geografiska data på Lantmäteriet, Geo SE, pågår arbete med utveckling av metadatahantering t ex för att möjliggöra publicering av data via katalogtjänster och öka spårbarheten i samband med leveranser till Försvarsmakten.

2.3.5 DGIWG

Inom den militära sektorn har DGIWG (*Defence Geospatial Information Working Group*) varit en aktör på geoområdet sedan bildandet 1983. Gruppens uppgift är att verka för standardisering inom geoområdet för försvarsorganisationerna i medlemsländerna och har därvidlag bl a tagit fram en egen metadataprofil som bygger på ISO 19115 och ISO 19139. Denna profil är tänkt att utgöra ett standardiserat gränssnitt till DGIWGs dataöverföringsstandard DIGEST (*Digital Geographic Information Exchange Standard*), som funnits i ett par decennier. Syftet är att DIGEST-data (t ex multinationellt samproducerade vektor- och rasterdatabaser) ska kunna delas mellan olika aktörer i infrastruktur som bygger på dessa ISO-standarder.

2.3.6 NATO

Inom NATO finns överenskommelser om standarder, så kallade *Standardization Agreements* (STANAGs) för att underlätta samarbete mellan NATO-länder. På geoområdet finns t ex sedan 1998 NATO STANAG 7074. Denna NATO-standard är detsamma som DIGEST, dvs NATO valde att adoptera DGIWGs standard.

Sedan 2006 har NATO utvecklat en infrastruktur för distribution av geodata – The NATO Core GIS⁸ – som svar på ett ökat behov att kunna förse olika militära enheter med aktuella geodata. Data görs tillgängliga mha standardiserade OGC-tjänster (WMS, WFS, WCS) så att i princip vilket system som helst med uppkoppling till NATO-nätverket ska kunna tillgodogöra sig data, förutsatt att det stödjer sagda tjänster. Ett utkast till metadataprofil för NATO Core GIS presenterade i december 2009 och baserat på denna har ett förslag till en XML-specifikation för en ny STANAG (2586) tagits fram.

2.3.7 USA

Efter 11 sep-attackerna i USA tog arbetet med interoperabilitet på geoområdet i USA ordentlig fart. NGA (*National Geospatial-Intelligence Agency*) bildade t ex ett centrum för att utveckla och koordinera civil-militärt standardiseringsarbete i USA, NCGIS (*National Center for Geospatial Intelligence Standards*). År 2005 bildades the *GEOINT Standards Working Group*⁹ (GWG), ett forum vars syfte är att driva på i frågor kring standardisering inom GEOINT-området. GWG fungerar som ett rådgivande organ till NGA, såtillvida att det är inom GWG konsensus skapas mellan olika aktörer (US DoD, CIA, NSA, OGC, ANSI, DGIWG, DARPA, ...) som sedan blir rekommendationer till NGA. GWG leds av NCGIS.

Konkreta resultat från detta omfattande multidisciplinära arbete är tänkt att förkroppsligas i form av NSG (*National System for Geospatial-Intelligence*) som syftar till att integrera teknik, politik, förmågor och doktriner i ett system för att skapa och underhålla en gemensam lägesbild baserat på ett enhetligt sätt att hantera geodata. Det sätt att specificera och koda geodata som föreslås inom NSG kallas GSIP (*GEOINT Structure Implementation Profile*). GSIP består i själva verket av ett antal komponenter, bl a kataloger som specificerar vilken semantik som ska användas, applikationsscheman (hur semantiken ska kodas), filformat för data, etc.

USA har tillsammans med Kanada också tagit fram en egen profil av ISO 19115. FGDC (Federal Geographic Data Committee) ledde arbetet som har resulterat i *North American*

⁸ <http://www.esri.com/industries/defense/pdfs/nato.pdf>

⁹ http://www.gwg.nga.mil/documents/GWG_Charter.pdf

*Profile*¹⁰ (NAP) of ISO 19115:2003. Standarden fastställdes av ANSI (American National Standards Institute) år 2009.

¹⁰ <http://www.fgdc.gov/nap/metadata>

3 Metadata för 3D-produkter

I det här kapitlet tittar vi närmare på metadata för 3D-produkter. På geoområdet i stort, och inte minst på höjd/3D-området, växer flera användningsområden fram. På det militära området börjar flera tillämpningar ta klivet från demonstrationer av vad som är möjligt rent tekniskt, till att integreras i system av olika slag. Det kan handla om siktfältsberäkningar, navigering, positionering, terränganalyser, radioberäkningar och mycket annat.

Nya tillämpningar kräver en väl strukturerad hantering av höjd/3D-data så att man så långt det är möjligt kan garantera att man genomför analyser och fattar beslut på solida grunder. Ytterst kan säkerheten för egen trupp eller tredje man stå på spel, och då gäller det att man har god kontroll på de data man använder och det är här som metadata kommer in.

3.1 Hur ser det ut idag?

Leverantörer av geodata skickar idag oftast med någon form av beskrivning av data. Tyvärr tenderar de ofta att beskriva sina data på ett helt eget sätt, vilket gör det svårt – ibland rent av omöjligt – att söka rätt på, värdera och jämföra olika datamängder. Ibland krävs det detektivarbete för att förstå hur data ska tolkas eller ens vilket koordinatsystem de representeras i. Information om datakvalitet är ofta särskilt bristfälligt beskriven och redovisad. Dessutom levereras informationen ofta på en form som är ämnad att tolkas av människor, t ex text-, PDF- eller Excel-dokument.

På mjukvarusidan är stödet för standarder idag begränsat. Enligt en mjukvaruleverantörs egna uppgifter kanske ISOs geodatastandarder följs, men man måste titta noggrant för att se det finstilla i ett sådant påstående. Stödjer systemet *hela* standarden eller är det begränsat till endast de *obligatoriska* delarna? Att kunna exportera en metadatafil som är en giltig ISO 19139-fil är en sak, att kunna exportera en metadatafil med ”all tänkbar” information som en viss tillämpning kräver är en annan.

Sammantaget tycks detta peka på en situation där data allt som oftast strömlinjeformas och skräddarsys för att passa en redan identifierad målgrupp som förväntas ha god bakgrundskunskap om den aktuella datatypen, eller som vågar förlita sig på dataleverantörens ”track record” för att bedöma huruvida data är lämpliga att använda i en viss tillämpning.

3.2 Vilka metadata behövs för 3D-produkter?

I det här avsnittet tar vi upp några metadataaspekter som projektet funnit särskilt viktiga för höjd/3D-data.

3.2.1 Översiktlig beskrivning

En grundläggande komponent i metadata är en översiktlig beskrivning av dataresursen. Denna komponent är generisk såtillvida att den behövs oavsett om själva data representerar höjdvärden, vägnät, byggnader, satellitbilder eller något annat. Förutom uppgifter om vem som ansvarat för att ta fram metadata, när detta skett, om uppdateringar av metadata planeras, syftet med att ta fram datasetet, etc, bör metadata innehålla en generell beskrivning av den aktuella resursen. ISO 19115 är uttrycksfull på den punkten och ger bl a möjlighet att beskriva resursens dataformat (*MD Format*), information om hur datas spatiella representation (*MD SpatialRepresentation*), upplösning (*MD Resolution*), uppgifter om ev. språk som förekommer i data (*language*), etc. Detta ger förutsättningar för att kunna hantera resursen rent tekniskt. Uppgifter om restriktioner och rekommendationer angående användning av data (*MD Constraints*) kan också bakas in i metadata. ISO 19115 tillåter att man i klartext anger och motiverar begränsningar i

data som påverkar dess lämplighet i olika tillämpningar men även att man specificerar ev. legala, sektretess- och affärsmässiga restriktioner på användning av data (*MD_RestrictionCode*, *MD_ClassificationCode*).

Kommentar:

De i ISO 19115 specificerade alternativen för sådana restriktioner är relativt generella och för militära ändamål behöver man sannolikt gå längre än så. Att i klartext ange vad data får och inte får användas till lämnar utrymme för oklarheter och är dessutom svårt att tolka automatiskt. Ett alternativ är att istället definiera en ”kodlista” med materielsystem eller användningsfall och koppla informationen till denna. Det ger ökad tydlighet och minskar risken för missförstånd och godtycke.

3.2.2 Referenssystem

Metadata bör innehålla uppgifter om vilket referenssystem som data representeras i. I sammanhanget är EPSG-koder¹¹ lämpliga att använda när så är möjligt. I en sådan kod kan plan- och höjdsystem kombineras till en enhet (*CompoundCRS*). Ett exempel är EPSG:5845 som är en kombination av SWEREF99 TM + RH2000. Dessa koder täcker inte in precis alla geodetiska parametrar som använts runtom i världen, men alla de mest frekvent förekommande systemen finns med. Genom att använda sådana koder kan dels missförstånd undvikas, och dels ökar förutsättningarna för att hantera data automatiskt.

Kommentar:

I ISO 19115 kan en sådan EPSG-kod anges men referenssystem kan även beskrivas på sedvanligt vis med ellipsoid- och projektionsparametrar.

3.2.3 Typ av höjdangivelse

Endast en uppgift om att datasetet innehåller höjdinformation (*elevation*) är i regel inte tillräcklig. För att man ska kunna nyttja data på rätt sätt måste det framgå vad höjdvärdet avser. Är det är markens höjd, den högsta uppmätta höjden i varje position eller kanske rentav höjden på de objekt som står på marken? Detta är en alltför central uppgift för att förbigås eller lösas ”ad hoc” med hjälp av nyckelord eller kommentarer i fritext. En sådan variant lämnar nämligen ett för stort utrymme för olika tolkningar av begrepp. Inom geområdet råder viss begreppsförvirring och det är viktigt att alla talar samma språk. Förvirringen gäller vanligen förhållandet mellan begreppen Digital Terrain Model (DTM), Digital Surface Model (DSM) och Digital Elevation Model (DEM).

Det här projektet har gjort följande tolkning av begreppen:

- DTM, markmodell: den uppskattade höjden på markytan (ibland kallad ”*bare-earth model*”), vanligen framställd genom bearbetning av en markestimeringsalgoritm
- DSM, ytmodell: representerar den högsta höjden, inklusive alla objekt på marken (typiskt det högsta värdet i varje position efter att felaktiga mätpunkter tagits bort)
- DEM, höjdmodell: ett generiskt begrepp som i princip kan avse vilken typ av höjdmodell som helst. Om man t ex har mätt höjdvärden som man inte kan härleda till markytan eller objekt på denna, är det lämpligt att använda begreppet DEM. Ett exempel är höjdmätningar mha interferometrisk SAR (InSAR), t ex det rymdskyttebaserade SRTM-projektet, där det kan vara svårt att avgöra vad som genererat det signalsvar som höjdvärdet baseras på.

Kommentar:

Vad höjdvärdet avser bör alltså vara en obligatorisk uppgift, med ett antal vedertagna och väl definierade termer att välja bland. Det skulle underlätta avsevärt för värdering,

¹¹ <http://www.epsg-registry.org/>

katalogisering och sökning bland datasetet. I ISO 19115 finns kodlistan *MD_TopicCategoryCode* där ett element är "elevation". En utökning av listan, eller en ytterligare lista, vore ett sätt att undanröja missförstånd. Ett exempel där ISO 19115 utökats med nya kodlistor och nya element i existerande listor (dock inte för höjdtypeper) är *North American Profile (NAP) of ISO 19115:2003 Metadata*. Exempel på utökningar i NAP jämfört med ISO 19115 är kodlistor över resursens filformat, typ av datorgränssnitt som data distribueras genom (t ex webbtjänster) och språket i metadata (ett "inlån" från ISO 639-2). Något liknande kan även göras för beträffande en finare uppdelning av höjddatabegreppet (vilket alltså inte gjorts inom NAP).

3.2.4 Ursprung och processhistorik

Uppgifter om datas ursprung och de bearbetningssteg som lett fram till det aktuella datasetet är också relevanta, inte minst eftersom höjd/3D-data ofta produceras genom en serie olika förädlingssteg. Historiken kallas ibland för *lineage* och ges stort utrymme i ISO 19139. *LI_Lineage* används för att beskriva *data* och *händelser* som legat till grund för datasetets tillblivelse, genom *LI_Source* respektive *LI_ProcessStep*. Dels kan man i fritext beskriva varifrån data kommer, men man kan också använda olika former av referenser och exempelvis peka ut resurser som är tillgängliga på internet (*CI_OnlineResource*), hänvisa till den som är ansvarig för ett visst bearbetningssteg (*CI_ResponsibleParty*) informera om huruvida datasetet är en delmängd av ett aggregerat dataset (*CI_Series*), eller till en publikation som beskriver datasetet (*ISBN/ISSN*), och beskriva på vilken form ursprungsdata finns (*CI_PresentationFormCode*).

Kommentar:

Detaljerad information om varje bearbetningssteg är kanske inte något som alla användare av data efterfrågar eller kan ta till sig, men det är troligen högst relevant för personal som framställer följdprodukter av data och ansvarar för deras kvalitet. Vi bedömer att de uttrycksmöjligheter som finns i ISO 19115 i detta avseende räcker långt och att utmaningen snarare ligger i att få dataleverantörer att delge informationen.

3.2.5 Täckning

För att ange ett områdes geografiska utsträckning är idag det allra vanligaste sättet att ange utsträckningen för den minsta omskrivande rektangel vars sidor är parallella med koordinatsystemets axlar ("bounding box"). I flera, bl a militära, sammanhang ökar emellertid möjligheterna till begränsad och riktad datainsamling som inte genererar stora sjöar av data som låter sig beskrivas med rektanglar. Istället kan data t ex bestå av flera mindre oregelbundet formade "öar", följa en slingrig väg, kraftledning, nationsgräns eller något annat som accentuerar behovet av en noggrannare beskrivning av datas faktiska utbredning.

Kommentar:

Vi anser att områdets utsträckning bör anges i form av polygoner som i ISO 19115 motsvaras av *EX_BoundingPolygon*. Dessutom bör man kunna indikera huruvida polygoner innehåller data eller representerar "hål" där data saknas (ISO 19115: *extentTypeCode*).

3.2.6 Kvalitetsmått

Uppgifter om datas kvalitet är mycket centrala. De bestämmer indirekt datas pålitlighet i olika tillämpningar och indikerar hur väl det aktuella datasetet går att kombinera med andra datakällor för mer komplexa analyser. Vår bedömning är att det finns mycket att göra på det här området. Att t ex ange endast ett kvalitetsvärde ("noggrannhet") för ett höjddataset som spänner över tiotals kvadratkilometer och flera terrängtyper är ett mycket trubbigt mått; det finns i själva verket ett flertal relevanta kvalitetsmått som dessutom kan variera kraftigt mellan olika terrängtyper.

Olika tillämpningar ställer naturligtvis olika krav på data och genom flera, mer detaljerade, kvalitetsmått får användare bättre stöd i sina analyser och bedömningar. ISO 19115 ger möjlighet att ange flera olika kvalitetsmått, med förklaringar och referenser till publikationer där dessa definieras eller beskrivs ytterligare. Både absolut och relativ noggrannhet kan t ex anges genom elementen *DQ_AbsoluteExternalPositionalAccuracy* och *DQ_RelativeInternalPositionalAccuracy* och för griddade data (på rasterform) kan *DQ_GriddedDataPositionalAccuracy* användas. En svårighet är att standarden är generell skriven för att täcka in alla former av geografiska data och att det därför kan vara svårt att direkt tolka den ur ett höjddataperspektiv. Exempelvis finns inget fördefinierat sätt att specificera noggrannheter i plan respektive höjd, utan detta får istället hanteras genom att ge måttet ett lämpligt namn, beskriva det eller hänvisa till en referens.

Det finns ett antal flitigt förekommande termer som brukar användas för plan- och höjdnoggrannhet. För plannoggrannhet är *Circular Error* (CE) ett vanligt mått. Ett CE90 om 1 m betyder att 90 % av kontrollpunkterna har befunnits ligga inom en meter från det verkliga värdet, men det säger inte någonting närmare om felets fördelning. För höjdnoggrannhet finns ett motsvarande mått i form av *Linear Error* (LE). Ett LE90 på 1 m kan alltså tolkas som att ett uppmätt höjdvärde med 90 % sannolikhet ligger inom en meter från rätt värde. Stöter man på CE och LE kan man alltså utgå från att det är frågan om noggrannhet i plan respektive höjd. En vanlig kvalitetsuppgift för (mark)höjdmodeller är det s k RMSE (Root Mean Square Error):

$$RMSE = \sqrt{\frac{\sum_{i=1}^N (\hat{z}_i - z_i)^2}{N}},$$

där $\hat{z}_i$ är den skattade markhöjden, z_i är det verkliga (noggrant inmätta) höjden och N är antalet kontrollpunkter. Den kvadratiske termen gör att redan av ett fåtal avvikande punkter får stort genomslag i RMSE-värdet även om majoriteten av värdena endast har små fel.

Uppenbarligen är ett enda kvalitetsvärde otillräckligt för att beskriva hur felen varierar, t ex som funktion av hur terrängen ser ut. På senare år har mycket forskning bedrivits om hur olika terrägegenskaper påverkar noggrannheten i höjdmätningar. En bra översikt och diskussion kring detta finns i [1].

Baserat på den existerande kunskapen om förhållandet mellan mätfel och terräng skulle metadata kunna vara betydligt mer informativa än vad som är fallet idag. I standarden ISO 19115 finns möjligheter att till varje datakvalitetsuppgift knyta information om för vad uppgiften gäller, ett s k ”scope”, t ex huruvida den avser hela den aktuella dataresursen eller en delmängd av data med gemensamma egenskaper. Det är t ex känt [1] att höjdmätningars noggrannhet påverkas av markens lutning på så sätt att osäkerheten är mindre för små lutningar och ökar påtagligt för större lutningar. För NNH är felet för snäll” terräng ofta betydligt bättre än vad som specificerats (< 0,5 m.) medan det i vegetation kan vara betydligt sämre. NNH-metadata¹² redovisar t ex en noggrannhet så bra som några få centimeter, men med vetskap om att kvaliteten på höjdmätningarna påverkas av bl a lutningen är det svårt att veta hur felet varierar i branter. Observera också att i NNH-fallet är data förädlad mha en markestimeringsalgoritm, vars prestanda i olika terrängtyper också påverkar resultatet.

Man kan idag mha ISO 19115 också associera ett kvalitetsmått med en geografisk utsträckning i form av en polygon (avsnitt 3.2.5). Detta ger möjlighet att definiera områden med olika kvalitet, t ex med olika lutning, olika terrängtyper (vegetation, öppeterräng, etc.) eller olika ursprung (t ex om den aktuella resursen är framställd utifrån flera underlag med olika kvalitet).

En annan relevant kvalitetsaspekt är temanoggrannhet (*thematic accuracy*). Den beskriver noggrannheten för en klassificering av datasetet. En tillämpning är markklassificering, där

¹² NNH-metadata för ett område som projektet fått tillgång till under 2012.

man typiskt kan ange hur många pixlar av en viss typ som förväntas vara felklassificerade. På samma sätt kan man beskriva klassificeringsprestanda för en markestimeringsalgoritm och beskriva hur stor andel av laserpunkterna som förväntas vara korrekt klassificerade.

Kommentar:

Sammanfattningsvis kan sägas att det finns flera väldigt relevanta standardiserade kvalitetsmått, men att standarderna inte tar explicit hänsyn till de ytterligare aspekter på detta som höjddata medför och därför kan leda till visst godtycke när kvaliteten hos sådana data ska beskrivas.

3.2.7 Kvalitetsvärderingsmetod

Hur ett kvalitetsmått ska uttydas är en sak, hur dess värden bestämts är en annan. Har någon utfört kontrollmätningar för det aktuella datasetet och jämfört med geodetiskt inmätta ”facit”-punkter? Och hur många punkter i sådana fall? Är olika terrängtyper representerade bland dessa eller avser värdet noggrannhet för en särskild sorts terräng, t ex stora platta ytor? Kvalitetsvärderingsmetoden kan i ISO 19115 anges som antingen *directInternal* (endast själva datasetet har använts i utvärderingen), *directExternal* (referensdata har använts) och *indirect* (baserat på tidigare erfarenheter). Genom att komplettera med beskrivning i fritext och referenser till kvalitetsutvärderingsprocessen kan tillförlitligheten till data höjas.

Kommentar:

Vi anser att det rimligen bör finnas krav på redovisning av vilken metod som använts för att ta fram ett visst kvalitetsmått. Detta kan göras helt i linje med existerande standarder. Det ställer högre krav på dataleverantörer och minskar osäkerheter och ”rykten” inom området som kan fortläpa sig genom att en lösryckt siffra accepteras som universellt giltigt utan att det finns någon spårbarhet bakom.

3.2.8 Upplösning

Begreppet ”upplösning” är en annan vanlig källa till missförstånd. Det är därför viktigt att så noggrant som möjligt specificera detta begrepp (i metadata) för att undvika oklarheter. En aspekt på upplösning har att göra med mätavståndet mellan de ursprungliga mätpunkterna (i ”rådata”). För högupplöst flygfotografering kan det röra sig om pixlar som motsvarar några få cm i sida på marken. För laserskanning brukar man t ex ange antalet laserpulser/m² eller antalet registrerade mätpunkter/m².

En annan aspekt är avståndet mellan dataelementen (pixlarna, punkterna) i det färdiga höjddataset till vilket metadata hör. Detta kan mycket väl skilja sig från de ursprungliga data. För NNH levereras t ex en rasterprodukt motsvarande 2x2 m² per pixel, medan rådata håller en högre upplösning och dessutom varierar över landet.

En ofta minst lika relevant aspekt är emellertid den egentliga rumsliga upplösningen, dvs storleken på den minsta geometri som höjdmodellen kan avbilda. Detta är ett betydligt mer komplext mått än de övriga, eftersom det påverkas av alla ingående steg i datainsamlings- och förädlingskedjan. Optiska egenskaper (linsdistorsion, laserlobens bredd, optiska filter, mm), plattformsegenskaper och databearbetning är faktorer som att påverkar slutresultatet. Det är därför vanskligt att uppskatta detta från enklare faktorer som pixelstorlek, avståndsbrus eller genomsnittlig punkttäthet.

Ponera exempelvis att man använder höjddata för att detektera förändringar mellan två mätningar, gjorda med olika mätsystem. Utan information om datasetens faktiska rumsliga upplösning blir det sannolikt väldigt svårt att tolka resultatet och att dra säkra slutsatser om vad som är verkliga förändringar och vad som är resultatet av bristande upplösning (och andra faktorer också för den delen). Har man däremot uppgifter om att objekt mindre än 30 cm inte avbildas korrekt, eller kanske inte ens finns med, så kan – och ska – inte data användas för att försöka detektera så små förändringar.

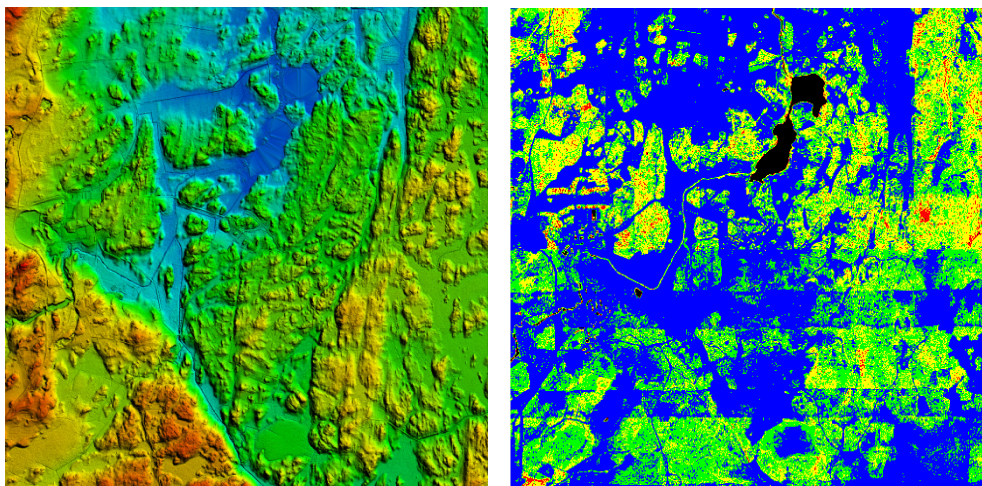
Kommentar:

I ISO 19115 finns inbyggt stöd för att uttrycka upplösningen i dataresursen, men där är fokus på upplösning i termer av avstånd mellan punkter eller pixlar, snarare än på hur små strukturer avbildas. Här vore det därför relevant att gå längre än standarden och explicit skilja på dessa olika tolkningar.

3.2.9 Aggregerade dataset

Dataset som på något sätt hör ihop kan, enligt ISO 19115, bilda ett *aggregerat dataset*. Man kan t ex ge en referens till en kompletterande dataresurs, definiera data som tillsammans med den aktuella produkten ingår i en högre nivåns resurs eller peka ut dataset som nyttjats i framställningen av den aktuella produkten. Vidare kan man t ex ange om dataseten härrör från samma mätkampanj, eller hänvisa till resurs som beskriver sensorn som använts i datainsamlingen, och så vidare.

För flygburen laserskanning är det idag vanligt att konstruera en bild som visar hur punkttätheten varierar över området. En sådan densitetsbild kan då göras till en egen dataresurs, med egna metadata, och sedan kopplas ihop med höjddata och bilda en aggregerad resurs. Se exempel på ett höjd- och densitetsbildpar i Figur 4.



Figur 4. Vänster: Färgkodad markhöjdmodell (NNH). Höger: Färgkodad densitetsbild (antalet markpunkter per ytenhet). Data har färgkodats för visualiseringsändamål för att belysa datavariationer inom området; de faktiska värdena är här inte viktiga.

Kommentar:

Detta sätt att koppla ihop resurser kan t ex användas för att knyta ihop ett dataset med någon form av kvalitetsinformation som inte låter sig beskrivas direkt i metadata. Ett exempel är *kvalitetsinformation i bildform*, som i Figur 4. Man kan på liknande sätt associera höjddata med en lutningsbild som räknats fram utifrån höjdbilden, eller relatera en markklassningsbild med ett flygfoto över samma område för att t ex möjliggöra tolkning av datas kvalitet som en funktion av terrängtyp.

4 Hur kan man skapa bättre metadata

4.1 Behovsinventering

Metadata kan ju beskrivas som ”data om data” och syftar då till att beskriva en datamängd på ett meningsfullt sätt och därmed ge förutsättningar för att data kan hanteras på ”rätt” sätt. Men vad innebär då ett meningsfullt sätt och vilka metadata stödjer detta.

Vilka metadata och kvalitén på dessa som behövs till en datamängd beror till stor del på de krav som tillämpningen ställer där dessa skall användas och bli meningsfulla. D v s det behövs kännedom om vilka tillämpningar/verksamheter som datamängden kommer att användas i så att en spårbar koppling mellan slutanvändarnas behov till de faktiska kraven på metadata kan erhållas.

Men det är inte bara rätt sorts metadata som behövs i en tillämpning för att man skall kunna uppnå rätt resultat, utan också rätt sorts geodata. Hur kan man då få fram vilka geodata och metadata som behövs i en tillämpning?

Inom systemutvecklingsområdet finns ett antal metoder som skulle kunna vara användbara för att få fram vilka krav som behöver uppfyllas i aktuell tillämpning. Dessa metoder används primärt för att identifiera vilka krav som ska specificera ett tekniskt system utifrån slutanvändarens behov. Men i det här fallet skulle det primära syftet inte vara att specificera ett tekniskt system, utan identifiera vilken geoinformation med tillhörande metadata som behövas i ett geodatasystem för aktuell tillämpning.

I bilaga 1 visas hur man kan arbeta utifrån systemutvecklingsmetoder, för att åstadkomma en spårbar koppling mellan slutanvändarnas behov till de faktiska kraven på geodata med tillhörande metadata.

En metod som har använts inom Försvarsmakten är GIBB (Geografisk informationsbehovsbeskrivning), som är en metod för att analysera och dokumentera en verksamhets behov av geografisk information. Metoden ägs av Lantmäteriet.

4.2 Krav på leverantörer

Metadataområdet idag är tämligen eftersatt. Leverantörer skickar med metadata efter eget gottfinnande och det tycks vara väldigt svårt att få standardiserade metadata, både avseende innehåll och form. Ofta krävs detektivarbete för att komma på vilket koordinatsystem som gäller, vilken upplösning data har, hur gamla data är, etc.

4.3 Standarder

Projektet anser att det är viktigt att ha metadata på en standardiserad form som gör att information går att extrahera automatiskt på ett robust och enhetligt sätt. Ett standardiserat och maskinellt läsbart metadataformat är en nyckel till att upprätta kataloger, sökbara databibliotek, minska behovet av manuell tolkning och omformattering av data och underlättar delande av data mellan system, organisationer och länder. Att få med stora programvaruleverantörer i detta är viktigt, eftersom det finns starka kommersiella intressen kopplade till proprietära lösningar som annars kan motverka den praktiska effekten av standardiserade lösningar.

Det är väldigt mycket som talar för att grunden till metadataarbetet bör utgöras av redan etablerade standarder inom området: ISO 19115/19139 (eller kommande uppdateringar av dessa). Emellertid kommer det sannolikt finnas delar av dessa standarder där vissa ändringar och tillägg bör göras för att tillmötesgå Försvarsmaktens behov. Detta kan t ex åstadkommas genom att skapa en ”FM-profil” av standarden. Profiler används ofta i GIS-världen för att effektivisera datahanteringen inom ett visst tillämpningsområde, t ex för att

begränsa en annars otympligt stor standard, definiera nya klasser utifrån existerande standardkomponenter eller skärpa minimikraven på vad metadata måste innehålla. Ett exempel i Sverige är den nationella metadataprofilen som innehåller de metadataelement som ansetts nödvändiga för publicering via Geodataportalen, inklusive de element som Inspiredirektivet kräver. En FM-profil skulle t ex kunna innehålla utökade ämnesordslistor (t ex typ av höjdangivelse), innehålla högre krav på kvalitetsrelaterad information (historik, kvalitetsvärderingsmetod, etc) eller utökade restriktioner angående datas användning.

4.4 Process för kvalitetsvärdering av 3D/höjddata

I avsnitt 3.2.7 diskuterades vikten av att beskriva data i termer av kvalitet och att dataleverantörer bör avkrävas uppgifter om hur kvalitetsmått tagits fram. För de fall då Försvarmakten får/köper/byter data från t ex utländska aktörer är sådan kvalitetsinformation viktig, och utgör kanske t o m skillnaden mellan att data används eller inte används.

I andra fall, t ex i samband med mer långsiktig utveckling eller testning av olika materielsystem och datakällor, där det finns möjlighet att genomföra mätningar i Sverige, är förutsättningarna annorlunda. Projektet anser därför att man bör kunna gå längre och utveckla en process för kvalitetsvärdering av 3D/höjddata för Försvarmakten som specificerar hur olika kvalitetsmått ska tas fram. Processen skulle kunna inkludera ett testområde med flera referensinmätta objekt och egenskaper, en serie tester och jämförelser mellan de aktuella data och referensdata och därtill ett antal väl definierade statistiska mått.

Testerna skulle omfatta olika terrängtyper som erfarenhetsmässigt är olika svåra att mäta in och som ger möjlighet att jämföra resultat för olika mättekniker (laser, fotogrammetri), system, flyghöjder, etc.

Syftet med detta skulle vara

- utvärdering och provning av system under utveckling
- att ge (nya, presumtiva) leverantörer möjlighet till opartisk kvalitetsvärdering av data
- att mäta hur kvaliteten påverkas av olika dataförädlingssteg och av kvaliteten hos navigationsdata (GPS, IMU)
- leveranskontroller av beställda data

Samma typ av tester går även att genomföra även för de fall där data av olika skäl inte går att få fram för ett sådant testområde, t ex mha noggrann inmätning mha totalstation, markstående laserskanner, el dyl.

5 Referensdata

Som bas för det experimentella arbetet i projektet har använts olika typer av höjd/3D-data från Strid-i-Bebyggelse-anläggningen på Markstridsskolan i Kvarn. Därvidlag har ett referensdataset etablerats, tänkt att kunna användas som testdata även i andra projekt.

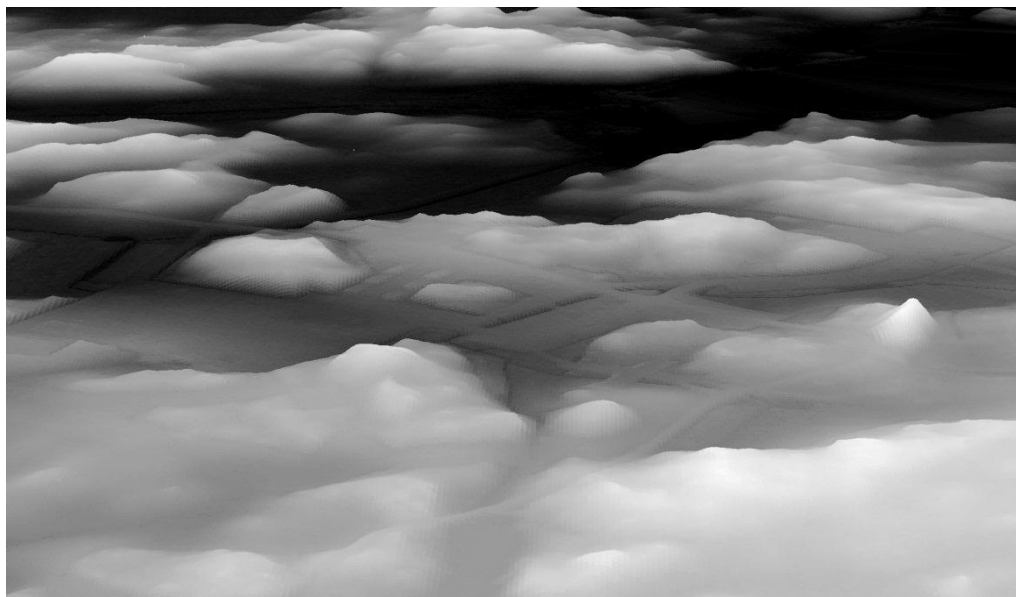
5.1 Data

Det finns flera olika sätt att representera 3D- eller höjddata. De vanligaste är *punkter*, *raster*, *höjdkurvor* och *TIN*. En representation av data i form av punkter ger användaren stor frihet att själv tolka punkterna och gruppera dem i ytor och objekt. I vissa fall kan det vara önskvärt, men det ställer samtidigt större krav på användarens erfarenhet av att tolka sådana data. Höjddata på rasterform är förhållandevis enkla att hantera maskinellt. Tack vare sin regelbundenhet är de väl lämpade för snabb databearbetning och olika typer av bildbehandlingsalgoritmer kan ofta användas mer eller mindre rakt av. En nackdel är att själva rasteringen innebär en förvanskning av data; värdet i en viss rastercell (pixel) är någon form av sammanvägning av värden från närliggande mätpunkter. En annan nackdel är att en höjdbild endast har ett höjdvärde per pixel och därmed inte utan speciallösningar kan representera överhäng, t ex broar. Ett TIN är ett nätverk av oregelbundet samplade data. En fördel med en sådan representation jämfört med ett raster är att man slipper den förvanskning som interpolation/rastering innebär. En annan fördel är att varierande ytor kan representeras med väldigt få trianglar så att merparten av data kan användas för områden där data varierar snabbare. En nackdel är att det ofta är mer komplicerat att maskinellt hantera TIN än data på rasterform.

Av praktiska skäl, såsom t ex datatillgång och mjukvaruverktyg för datahantering och visualisering, har inte alla datarepresentationstyper getts lika stort utrymme i projektet. Ur en metadatasynvinkel är emellertid frågor om själva datarepresentation av mindre betydelse eftersom metadata i sig inte är knutna till en specifik datarepresentation. Man kan mycket väl tänka sig flera dataset med olika representationer knutna till samma ursprungliga datamängd, så att användare/system kan arbeta med den representation som passar bäst givet en viss tillämpning. Det bör således i metadata anges hur data representeras.

5.1.1 NNH

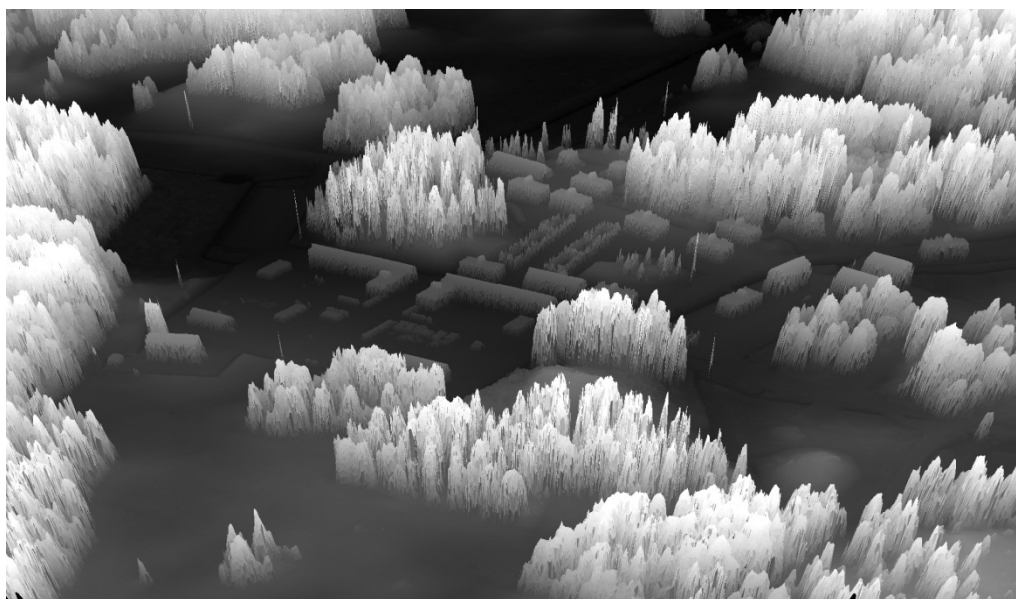
Lantmäteriet ansvarar för framtagandet av den Nya Nationella Höjdmodellen (NNH). Grundprodukten är ett raster med celler/pixlar om $2 \times 2 \text{ m}^2$ som representerar den estimerade markhöjden för respektive cell (Figur 5). Utöver denna finns också de data som legat till grund för rasterprodukten, nämligen ett punktmoln där varje punkt klassificerats som hörande till någon klass (t ex mark).



Figur 5. 3D-vy av NNH (Strid-i-bebyggelse-anläggningen på MSS Kvarn i mitten).

5.1.2 Högupplöst flygburen laserskanning

Projektet har även fått tillgång till högupplösta data från flygburen laserskanning över Strid-i-Bebyggelse-anläggningen (Figur 6). Mätningen genomfördes sommaren 2012 av företaget Sky Movies AB och håller en punkttäthet på ca 70-80 punkter/m². Efterbearbetningen av data gjordes av Foran Remote Sensing AB.



Figur 6. 3D-vy av en högupplöst höjdmödel från flygburen laserskanning (Strid-i-bebyggelse-anläggningen på MSS Kvarn i mitten).

5.1.3 R3DM

Ett alternativ till höjdmätning med laser teknik är att använda fotogrammetriska metoder. Nuförtiden finns automatiserade databearbetningsmetoder som gör det möjligt att framställa 3D/höjdmödel för stora områden utifrån kamerabilder. I Figur 7 visas ett exempel på detta i form av Saabs Rapid 3D Mapping-teknik.



Figur 7. Exempel på högupplösta 3D-data skapade med fotogrammetriska metoder (Saabs Rapid 3D Mapping). Bilden visar Strid-i-bebyggelse-anläggningen på MSS Kvarn.

5.1.4 Markbunden 3D Laserskanner

Strid-i-bebyggelse-anläggningen på MSS i Kvarn utanför Borensberg har blivit geometriskt inmätt med en 3D laserskanner, Laserskanner Riegl VZ-400. Riegl VZ-400 är en av de bästa laserskannrarna på marknaden avseende precision, noggrannhet och upplösning. Precisionen i varje inmätt laserpunkt är ca 3 mm vilket anses vara riktigt bra, avståndsnoggrannheten är ca 5 mm. För denna mätning användes i huvudsak en vinkelupplösning på 0.04 grader. Anläggningen mättes in med 75 fasta positioner väl utspridda i anläggningen. De flesta positionerna mättes från en plattform monterad på ett biltak (Figur 8) vilket gjorde det både lätt och snabbt att förflytta sig mellan positionerna. Några positioner mättes från hustak och torn i anläggningen. Från varje fast position samlades drygt 10 miljoner punkter in och totalt antal exteriört inmätta punkter av anläggningen torde vara uppåt 800 miljoner punkter.

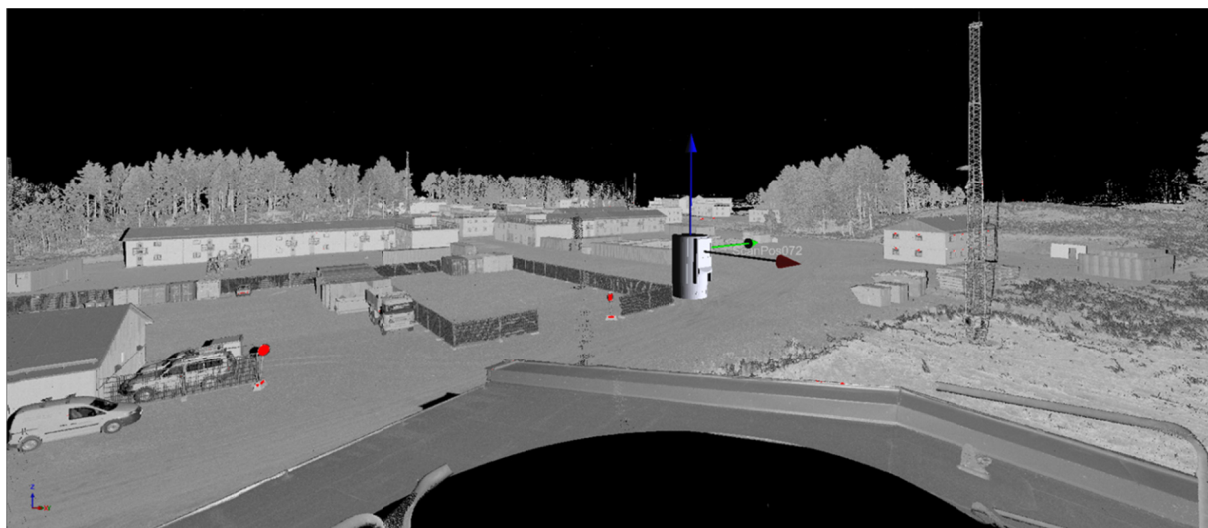
Från varje position har ett punktmoln samlats in bestående av lokala XYZ-koordinater. Varje position har sedan registrerats mot varandra i ett lokalt koordinatsystem. Till sist har hela det registrerade punktmolnet av anläggningen transformerats till globalt koordinatsystem (WGS84). Punktmolnet kan textureras med kalibrerad reflektans (Figur 9), dvs att korrigering görs för att kompensera att avståndet mellan de registrerade punkter ökar ju längre bort avståndet till det registrerade objektet är. Vid mätningarna togs även integrerade foton som kan användas för färgtexturering av punktmoln (Figur 10).

Hela det registrerade punktmolnet uppvisar en standardavvikelse på drygt 8 mm vilket får anses vara mycket bra.

Eftersom man från markbunden laserskanning erhåller en geometrisk avbildning med stor precision och punkttäthet, kan modellen användas för att verifiera andra 3D bildalstrande sensorsystem. Detta har varit huvudsyftet med denna geometriska mätning av anläggningen.



Figur 8. Riegl VZ-400 monterad på biltak. Från ett tidigare fältförsök i tyska Bonnland.



Figur 9. Punktmoln insamlat från byggnad 1 i Strid-i-bebyggelse-anläggningen på MSS Kvarn. Texturen på punkterna är kalibrerad reflektans.



Figur 10. Sydvästra området, här färgtexturerat. Infälld bild visar byggnad 1.

5.2 Metadata för referensdata

Som grund för det praktiska arbetet med metadata har projektet valt standarden ISO 19115:2003 beträffande innehållet i metadata. För kodning av metadata har ISO 19139 använts. Dessa val faller sig naturligt eftersom dessa båda standarder utgör grunden i mycket av den utveckling som sker nationellt och internationellt (avsnitt 2.3). Nya och uppdaterade standarder är visserligen under utarbetning, men har ännu inte accepterats.

För att implementera en metadatafil behövs i praktiken hjälpmedel som gör att man slipper skriva filen helt från början, som ger stöd angående hur filen ska byggas upp och som kan validera att den resulterande filen är korrekt. I projektet har två verktyg för detta studerats: metadataeditorn i ArcGIS 10.0 (från www.esri.se) och MetadataEditor 3.0 (från www.geodata.se). Båda verktygen har begränsningar som innebär att mycket av innehållet i ISO 19115 inte stöds. Obligatoriska uppgifter i metadata täcks in, men editorerna erbjuder ringa stöd för användare som vill gå längre och utnyttja den fulla potentialen hos ISO 19115/19139. Exempel på sådant som saknas är möjlighet att lägga till flera olika sorters datakvalitetsrapporter, ange områdets geografiska utsträckning som något annat än en "bounding box" (t ex polygoner), koppla kvalitetsmått mot geografiska områden, etc. Arbetsgången blev därför att först göra en metadatafil så långt som möjligt med något av verktyget och därefter komplettera filen med ytterligare information i en vanlig texteditor.

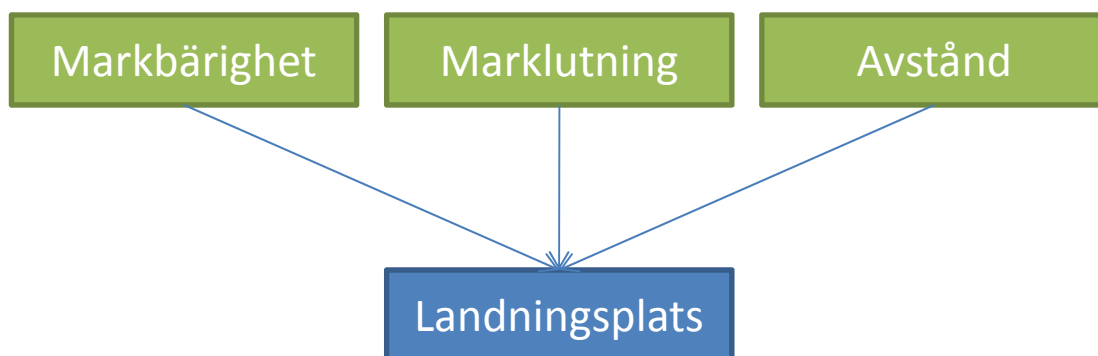
6 Testfall/scenario

För att åskådliggöra nyttan med geoinformation och metadata som är kopplad till geoinformationen, har ett scenario tagits fram för att mera konkret visa på nyttan med metadata. Scenariot (fiktivt) gäller framtagande av underlag för en potentiell helikopterlandningsplats utifrån de geodata som finns och de parametrar (fiktiva) som gäller för en helikopterlandningsplats. De ingående geodata utgör ett exempel på heterogena data med olika härkomst och upplösning.

Med hjälp av metadata kan geodata (inkluderande 3D data) tillsammans med övriga data, analyseras för att få fram lämpliga landningsplatser och för att bedöma möjliga felfortplantningar i geodata som kan påverka resultatet. Resultatet visualiseras sedan så att en operatör lätt skall kunna bedöma de föreslagna områdena.

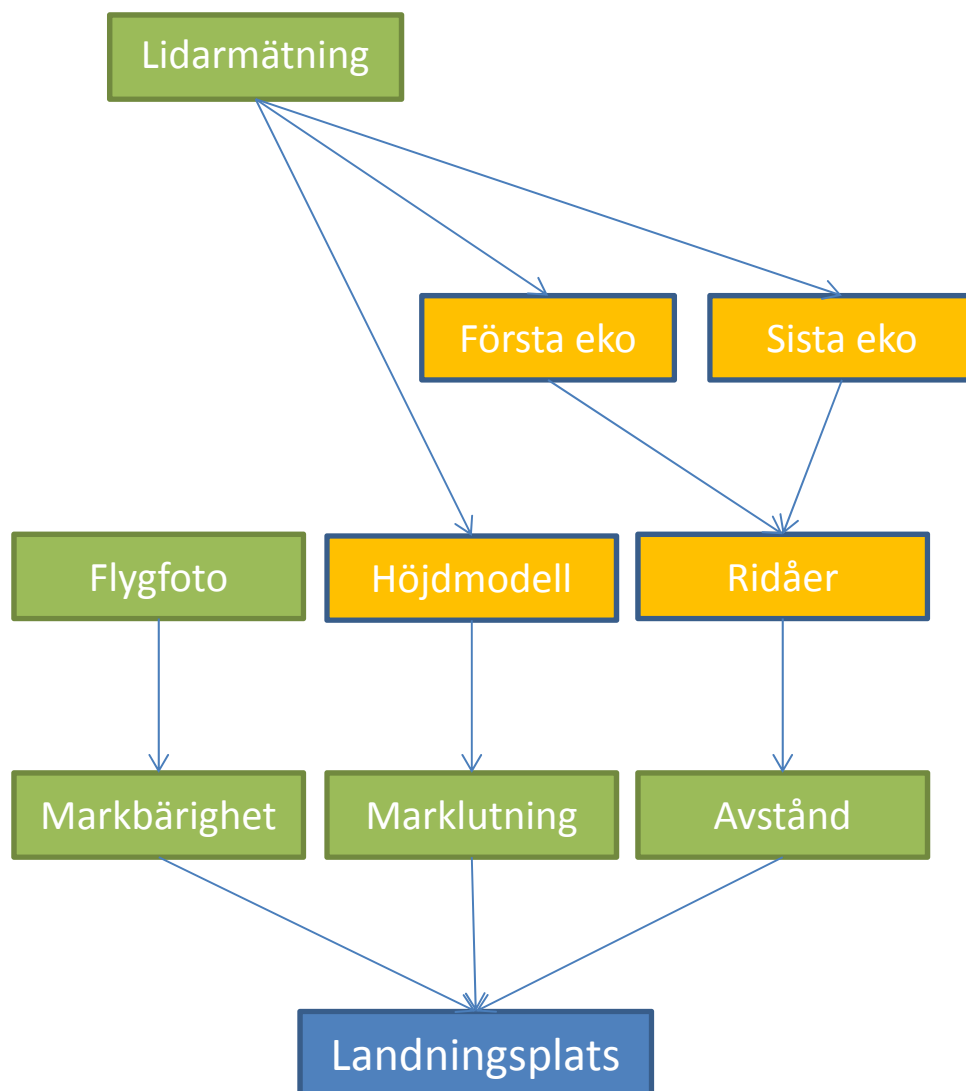
6.1 Beskrivning av scenariot

Nedan följer ett mycket förenklat exempel hur man kan resonera kring kartprodukter, metadata, kvalitet och analys. Vi tänker oss att det finns ett reglemente som beskriver hur en helikopter av en viss typ får agera under vissa givna omständigheter i ett givet operationsområde. Utifrån detta reglemente tas ett antal produktspecifikationer fram över olika kartprodukter. I specifikationerna står vilken information som skall ingå i respektive kartprodukt och vilken kvalitet produkten och de bakomliggande produkterna måste ha. Med bakomliggande produkter avses sådana kartprodukter som används i en eller annan form för att skapa den slutliga produkten. En sådan slutlig kartprodukt antas beskriva potentiella landningsplatser och har de bakomliggande produkterna markbärighet, marklutning samt avstånd till horisontrida så som skog, hus eller liknande. Denna tänkta slutprodukt får tjäna som exempel (Figur 11).



Figur 11. Produktbeskrivning. Slutprodukten Landningsplats härleds från de bakomliggande produkterna Markbärighet, Marklutning samt Avstånd till horisontrida. Pilarnas anslutningspunkt kan ses som en algoritm eller metod som går från de bakomliggande produkterna till slutprodukten.

Vidare så antar vi att de bakomliggande produkterna inte finns tillgängliga utan i sin tur måste härledas från andra kartprodukter. Detta skapar en hierarki av produkter, ett exempel på en sådan hierarki ses i Figur 12. Vi går senare in i detalj på denna hierarki.



Figur 12. Det kan krävas en hierarki av produkter för att få fram slutprodukten "Landningsplats", markerad med blått. De gul markerade produkterna är härledda produkter; 'Första eko' och 'Sista eko' är en rasterytmodell skapad från första respektive sista laserreturen. 'Första eko' mäter t ex i en skog höjden på trädtopparna medan 'Sista eko' ofta mäter höjden till mark eller tät undervegetation. Skillnaden mellan dessa benämns här 'Ridåer'. Varje laserpuls innehåller således från marken uppstickande halvtransparenta föremål och objekt. Höjdmodell är en ytmodell där byggnader, vegetation och andra uppstickande föremål filtrerats bort. De grönmarkerade produkterna är produkter som i vårt exempel återfanns på hyllan. Varje anslutningspunkt kan ses som en algoritm eller metod för att gå från en eller flera bakomliggande produkter till en slutprodukt.

6.1.1 Reglemente

Vi tänker oss att det finns ett reglemente som föreskriver under vilka förhållanden en helikopter får landa. Exempel på sådana förhållanden kan vara maximal marklutning, förekomsten av uppstickande föremål, avstånd till horisontridåer så som hus, skog och liknande, markbärighet, maximalt tillåtna vindstyrkor, risk för påskjutning samt avstånd till annan infrastruktur. Vi tänker oss också att det finns tre klasser, grön, gul och röd som motsvarar olika risknivåer för respektive förhållande. Grönt motsvarar att förhållandena håller sig inom acceptabla gränser, gult innebär att förhållandena är inom tolerabla gränser men att försiktighet skall iakttas samt rött att förhållandena är för riskfyllda och att det därför inte är tillåtet att landa. Projektet syftar dock inte till att ta fram en metod för att identifiera potentiella helikopterlandningsplatser utan till att påvisa nyttan med metadata. Vi tillåter oss därför att förenkla problematiken så att detta kan belysas. I vårt exempel begränsar vi oss till en kartprodukt som innehåller klassad marklutning, klassat avstånd till

horisontridå samt klassad markbärighet. Tabell 1 sammanfattar vilka krav och klasser som har använts i exemplet.

Tabell 1. Utdrag ur ett tänkt reglemente för landning av helikopter av en viss typ. Här ställs krav på bland annat maximal marklutning, förekomsten av uppstickande föremål, minimalt avstånd till horisontridåer så som hus, skog och liknande, markbärighet, osv

Lager	Grön	Gul	Röd
Markbärighet	Hård yta	Medelhård yta	Mjuk yta
Avstånd till horisontridå	Mer än 20 meter	10 – 20 meter	Mindre än 10 meter
Marklutning	Mindre än 3 grader	3 – 6 grader	Mer än 6 grader

6.1.2 Produktspecifikation

Utifrån reglementet tas en eller flera produktspecifikationer fram för helikopterlandningsplatser. Reglementet samt helikoptertyp ställer krav på vilken kvalitet indata måste ha. Skall marklutningen skattas över ett område motsvarande helikopterns utbredning med någon grads noggrannhet måste självfallet lutningslagret ha en noggrannhet väl inom dessa gränser. I vårt exempel tänker vi oss att en sådan produktspecifikation föreskriver att vi skall hitta områden med gynnsamma förhållanden för landningsplatser samt identifiera områden där landning är olämplig med hänsyn tagen till markbärighet, avstånd till horisontridåer samt marklutning. Förhoppningsvis motsvarar våra krav till stora delar en eller flera redan befintliga standardprodukter där rutiner och metoder redan är framtagna av någon geofunktion i organisationen, annars krävs det en del extra arbete. Man får i sådana fall fundera över om produktsortimentet skall utökas eller om de uppkomna kraven var en enstaka händelse. Tabell 2 sammanfattar de krav som ställs på de olika lagren i vårt exempel. Kraven är i detta fall fabricerade men i ett verkligt fall framtagna av expertis.

Tabell 2. Del av produktspecifikation för helikopterlandningsplatser. Fabricerade krav på ingående lager. En verklig produktspecifikation är väsentligt mer omfattande. Huruvida ett datalager uppfyller dessa krav bör framgå från metadata.

Lager	Spatiell noggrannhet	Värdenoggrannhet
Markbärighet	1 meter	> 80 %
Avstånd till ridå	1 meter	< 1 meter
Marklutning	1 meter	< 0,25 grader
Samtliga	1 meter inbördes	
Resultat	1 meter inbördes, 4 meter absolut	

6.1.3 Metod

För att påvisa nyttan med metadata tas de olika lagren som anges i produktspecifikationen fram med så bra underlag som finns tillgängligt. I detta fall har en laserskannad höjdmmodell använts tillsammans med ett färgflygfoto insamlat vid samma tillfälle. Lasermodellen har en punkttäthet kring 60 – 70 punkter per kvadratmeter och flygfotot en upplösning på ca 10 cm. Datat samlades in under april 2012. Lasermodellen har omvandlats till två rastermodeller med 0,5 meters upplösning i plan; den första modellen utifrån den från första laserreturen och den andra modellen från sista laserreturen. Dessa modeller har sedan i sin tur använts för att skatta marklutning och avstånd till horisontridåer. Även flygfotot har samplats om, i detta fall till 2 meters upplösning, denna

har i sin tur legat till grund för markklassningen. Dessa omsamlingar motsvarar ungefär den typ av data som finns tillgängligt idag eller relativt enkelt kan samlas in. Redan vid dessa omsamlingar kan man göra en analys för brus- och felfortplantning men det tjänar i vårt fall inget syfte utan de resulterande lagren utgör i vårt exempel referensdata och anses fria från brus samt placerade på exakt korrekt plats i världen. Vi låter sedan dessa data genomgå olika typer av försämringar och beskriver dessa i metadata och visar sedan hur man med hjälp av metadata kan analysera datat. Denna analys kan göras på en mycket avancerad nivå, men vi har här valt en nivå lämplig för att belysa konceptet och inte de ingående delmomenten i detalj. Området som används i detta exempel är 'Strid i bebyggelse'-anläggningen på MSS Kvarn (Figur 13).



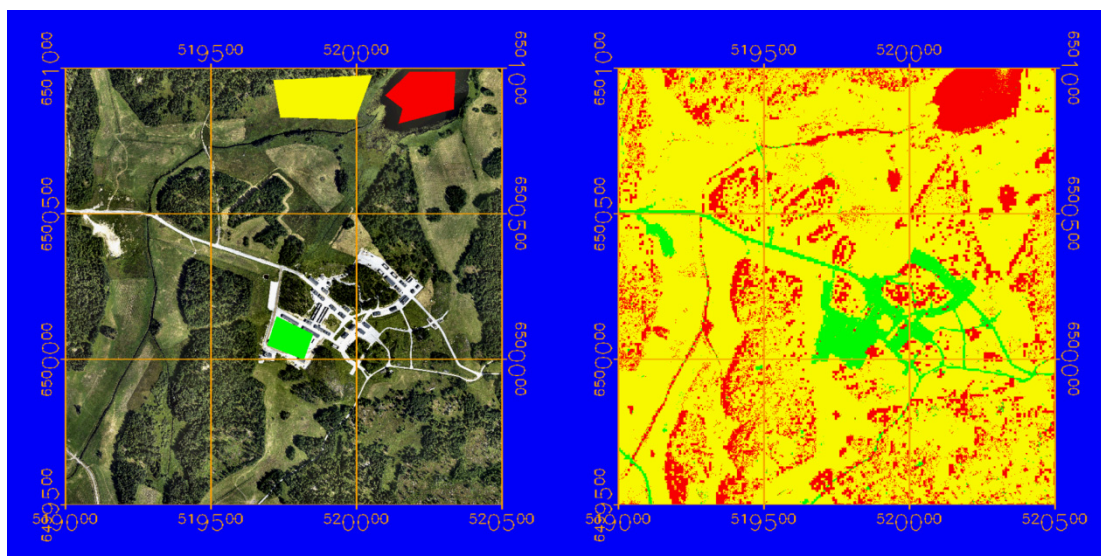
Figur 13. Färgflygfoto över Strid-i-bebyggelse-anläggningen i Kvarn. Ett något mindre område har använts för själva höjdmodellen.

6.1.4 Ingående datalager

Markbärighet har klassats utifrån ett digitalt ortorektifierat flygfoto. Marklutning och avstånd till horisontridaer har skattats från en laserskannad höjdmodell insamlad vid samma tillfälle som flygfotot togs.

6.1.4.1 Markbärighet

Markbärigheten har skattats genom att klassa ett färgflygfoto med 2 m upplösning. En s.k. *Support Vector Machine* [2] har fått träningsdata från tre områden som representerar vatten, vegetation samt urban terräng (Figur 14). Som man kan se i bilderna är det inte en speciellt god klassning. Algoritmen har svårt att skilja på vatten och skugga men å andra sidan fångas både vattendrag och vägar upp relativt bra. För vårt exempel duger det dock. Vi anser därefter att områden som klassats som urban terräng är lämpliga (grön) landningsområden, områden som klassats som vegetation är mindre bra (gul) områden samt områden som klassats som vatten som olämpliga (röd) områden.



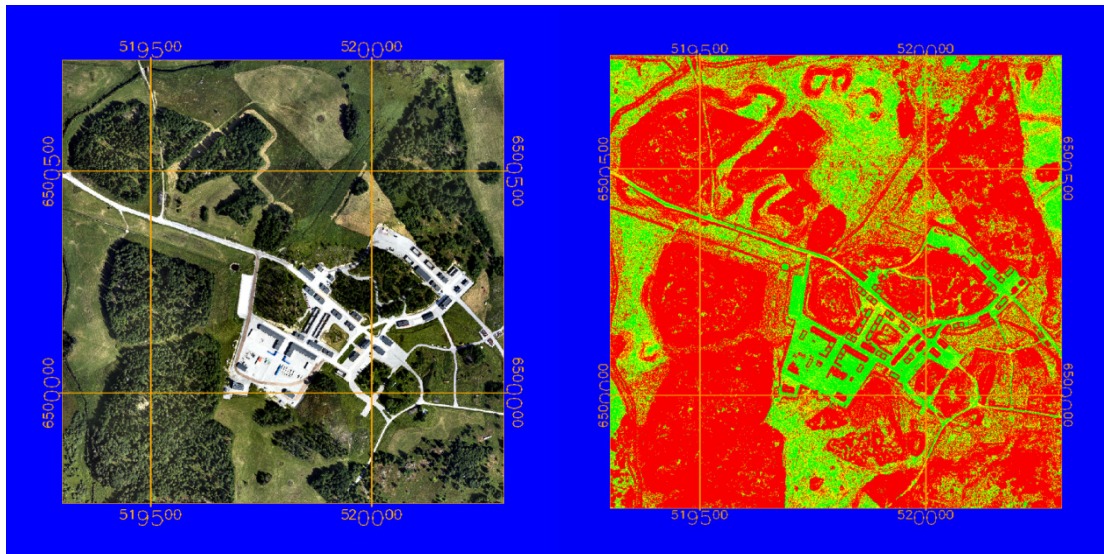
Figur 14. Markklassning. Bilden till höger visar de områden som använts vid träningen av en *support vector machine*. Till höger ses resultatet där rött motsvarar vatten, gult vegetation samt grönt urban terräng. Man kan konstatera att algoritmen har svårt att skilja på vatten och skugga varför en del vegetation klassas som vatten. Skuggorna i vegetationen kommer huvudsakligen från skogsridåer.

6.1.4.2 Lutning

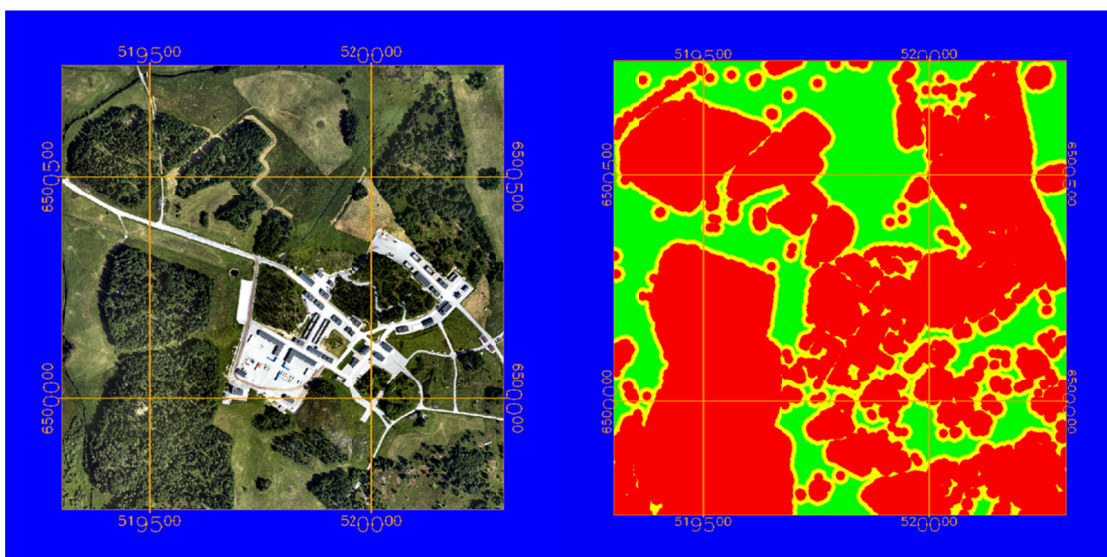
Lutningen (Figur 15) har skattats utifrån höjdmodellen med hjälp av mjukvaran IDL/ENVI's inbyggda algoritm för lutningsskattningar. Dokumentationen för IDL/ENVI redovisar inte exakt hur lutningen skattas men får här ändå tjäna sitt syfte som exempel.

6.1.4.3 Avstånd till horisontridå

Även avståndet till horisontridå (Figur 16) har beräknats utifrån en laserskannad höjdmodell. Horisontridå har här definierats som föremål som sticker upp mer än 1 meter. Skillnaden mellan första och sista ekot från laserskanningen har fått tjäna som underlag för att hitta uppstickande föremål då första ekot ofta reflekteras från trädtopparna medan sista ekot reflekteras från marken eller tät undervegetation. Givetvis kan man lägga in andra källor så som kraftledning och kända master men så har inte gjorts i detta exempel. Skillnadsbilden har därefter trösklats vid en höjd av en meter varpå en avståndstransform har beräknats och sedan klassats. Lämpliga (grön) landningsområden har ett avstånd större än 20 m till närmsta horisontridå, mindre lämpliga (gul) områden har ett avstånd från 20 till 10 m, olämpliga (röd) områden har ett avstånd på mindre än 10 m.



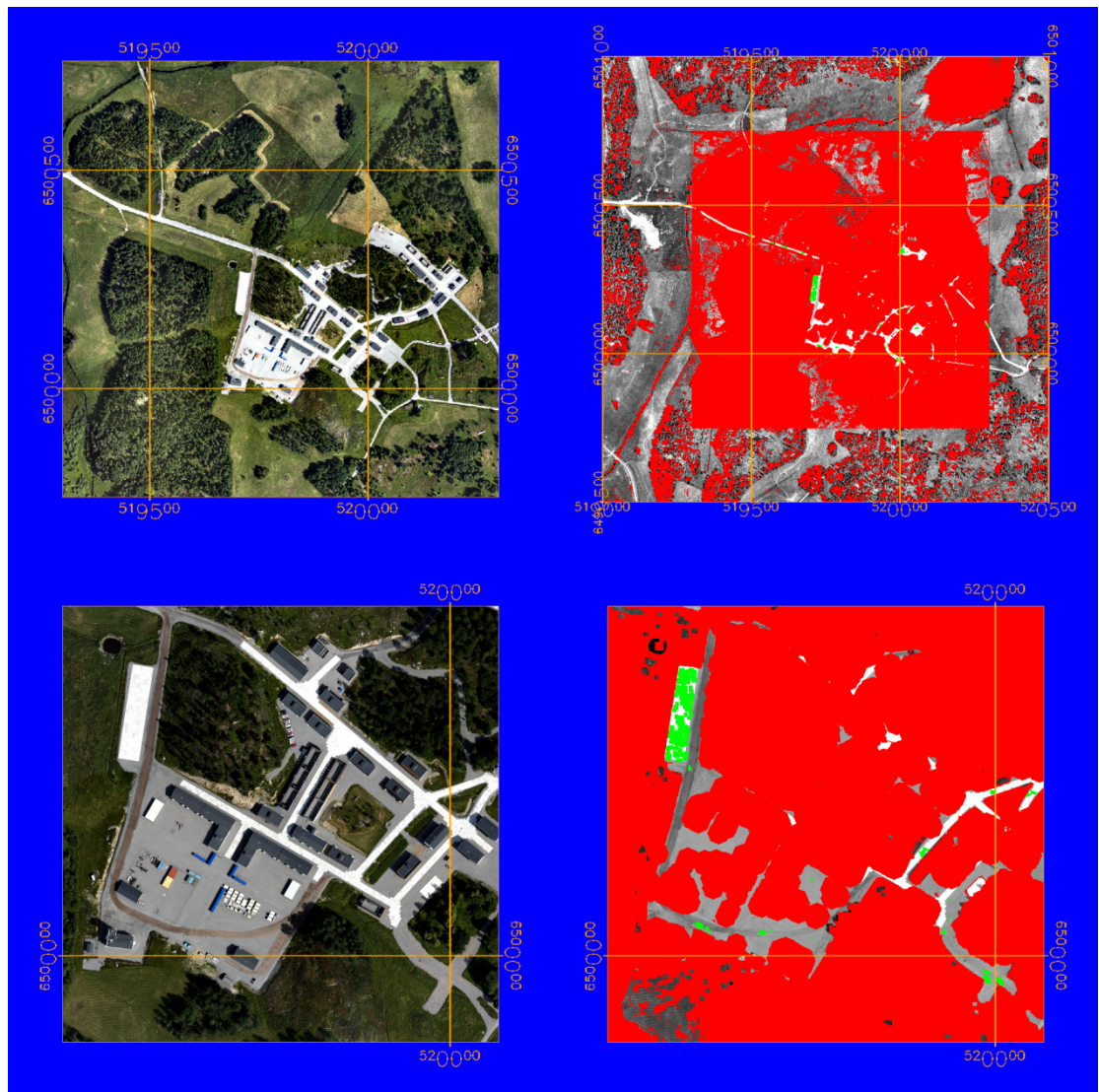
Figur 15. Lutning. Till vänster visas ett flygfoto över området som orientering. Till höger visas en klassad lutningsbild där rött motsvarar en lutning på mer än 6 grader, gult en lutning mellan 3 och 6 grader och grönt en lutning mindre än 3 grader.



Figur 16. Avstånd till horisontrida. Till vänster återfinnes ett flygfoto för orientering. Till höger den klassade avståndsbilden. Rött motsvarar ett avstånd mindre än 10 m, gult ett avstånd mellan 10 och 20 m samt grönt ett avstånd mer än 20 m.

6.1.4.4 Exempeldata

Kombinerar man dessa lager enligt reglerna att gröna områden är sådana områden där samtliga ingående lager har grön klassning och att röda områden är områden där minst ett lager har röd klassning fås resultatet enligt Figur 17. Vi låter detta resultat tjäna som facit eller referensdata i detta exempel för att utvärdera hur metadata påverkar analysen.



Figur 17. Resultat. Längst upp till vänster återfinnes ett flygfoto för orientering. Längst upp till höger ses resultatet. Gröna områden är områden där förhållandena är acceptabla i samtliga lager. Röda områden är områden där ett eller flera lager har indikerat att det inte går att landa. Den kvadratiska formen som kan ses i bilden kommer sig av att flygfotot som använts som underlag för markklassning skiljer sig i storlek jämfört med höjdmodellen som tjänat som underlag till de andra lagren. Nedre raden är ett uppförstorat område av ovanstående bilder. Ofärgade områden är områden där ett eller flera datalager indikerat att det är mindre bra att landa men inget lager indikera att det är olämpligt att landa.

6.1.5 Felfortplantning och metadata

Data som samlas in är alltid behäftat med en viss mängd brus. Bruset kan ha många källor och härstammar nästan alltid från den termiska bakgrundsstrålningen som ger störningar på detektorer och i förstärkare; andra typer av störningar kan förekomma som har andra fysikaliska förklaringar, t ex klotter i radarfallet, pulsbreddning i lidar-fallet eller atmosfäriska störningar i GPS-fallet. Dessa bakas ofta, för respektive fall, (något felaktigt) samman i begreppet brus och ger ett mått på vilken avvikelse mätningen ger från det sanna värdet. Sker en förädling av data som i fallet då lutningen skattades utifrån en höjdmodell måste hänsyn tas till detta och hur detta brus fortplantas genom de algoritmer man använder. Typiskt brus för en lidarinmätt höjdmodell ligger, på en plan hårdgjord yta, på någon centimeter. I gräs och annan vegetation får man en s.k. pulsbreddning vilket leder till betydligt högre brus. I vårt exempel har vi antagit att bruset är konstant oavsett terrängtyp. Vi använder också begreppet *standardavvikelse*, vilket är ett mått på

spridningen av bruset, för att kvantifiera bruset. För noggrann analys måste man även koppla en statistisk fördelning till spridningsmättet, vissa fördelningar kräver även ytterligare spridningsmått för att beskriva bruset. Denna information går sällan att utläsa direkt ur data utan måste anges som metadata kopplat till respektive datalager. I vårt exempel bryr vi oss dock inte om vilken fördelning bruset har utan nöjer oss med att beskriva det i termer om *standardavvikelser*. Vill man gå från spridningsmått till sannolikheter och uttalanden med konfidens måste man dock ta hänsyn till fördelningen, antingen teoretiskt eller med hjälp av simuleringar.

Ett sätt att skatta lutningen α utifrån en höjdmodell är genom uttrycket

$$\alpha = \sqrt{\left(\frac{dh}{dx}\right)^2 + \left(\frac{dh}{dy}\right)^2}$$

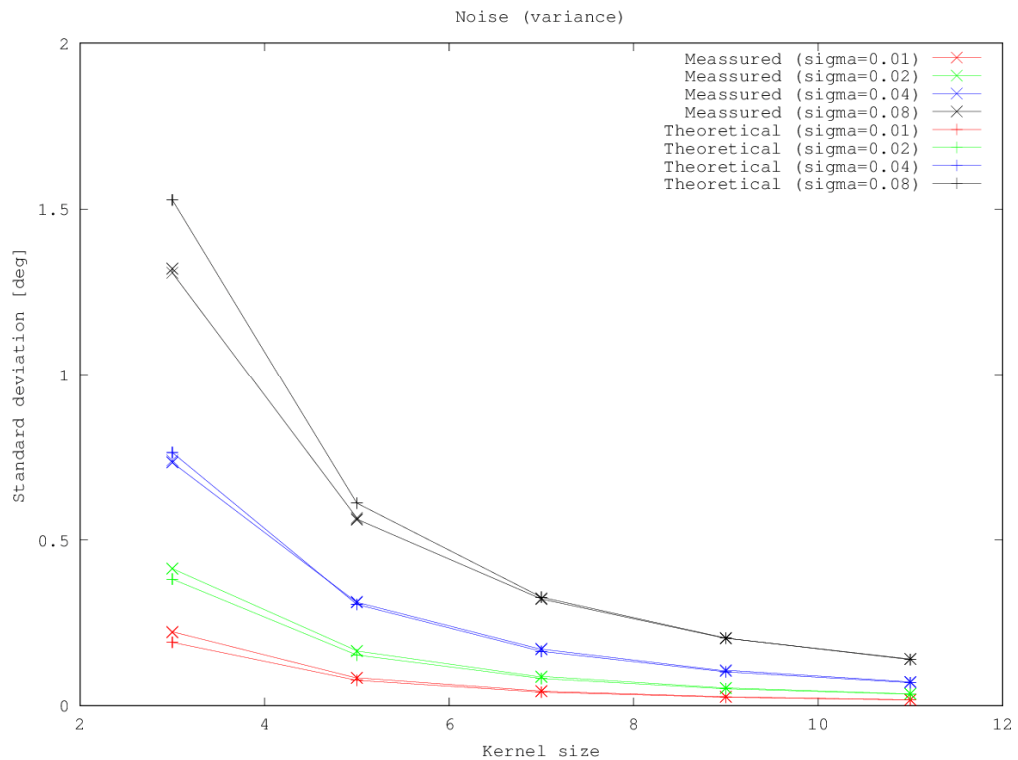
där $h(x, y)$ är höjdmodellen. Låt $\hat{h} = h + \tilde{h}$ där $\tilde{h}$ är normalfördelat vitt brus med standardavvikelsen σ . För ett raster kan man då visa att bruset $\tilde{\alpha}$ för små vinklar kan uttryckas som

$$\tilde{\alpha} \approx \frac{16\sigma^2}{D^2(D+1)^2}$$

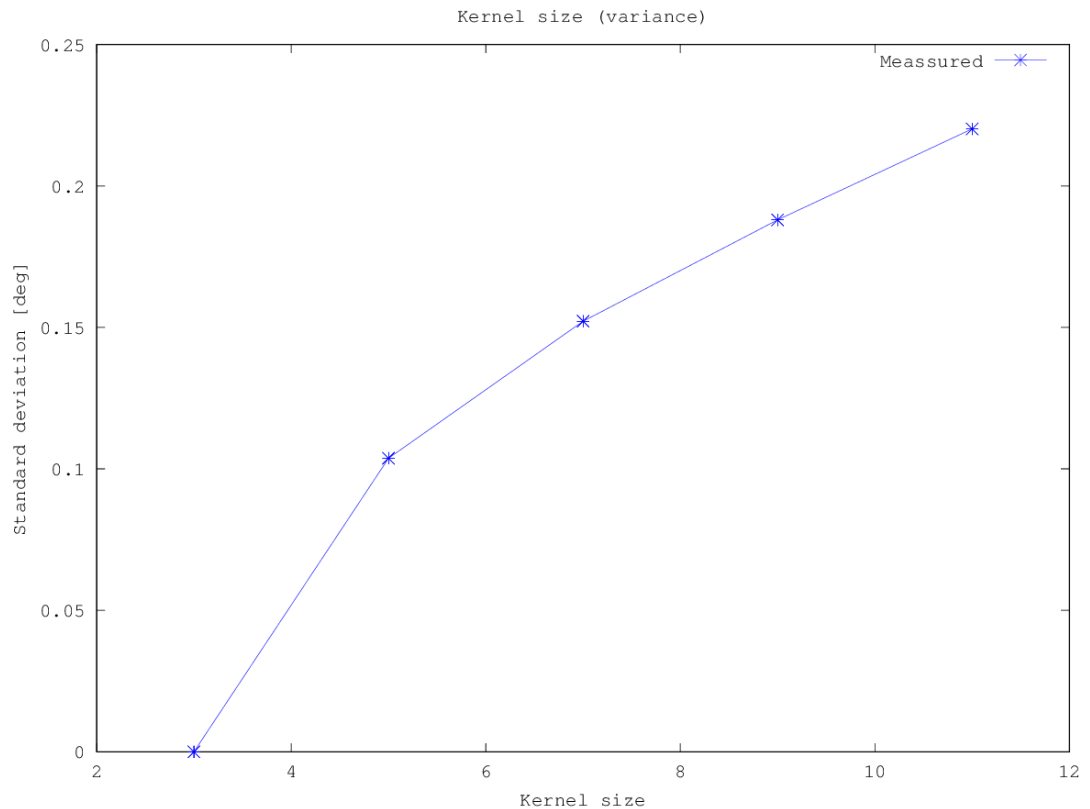
där D är storleken på de områden som använts för att skatta de lokala lutningarna i x- och y-led. Tack vare detta uttryck kan man bedöma om den underliggande höjdmodellen duger för att skatta lutningen utifrån i metadata angivet brus. Detta samband är inte alltid lätt att ta fram och det går inte alltid att göra en matematisk analys. Då kan man ta till simuleringsbaserade metoder istället. Figur 18 visar hur resultatet från en sådan simulering kan se ut. En helt artificiell höjdmodell har genererats varpå algoritmen för att skatta lutningen har applicerats för olika storlekar på områden varifrån de lokala lutningarna i x- och y-led skattats. I såväl det teoretiska som det simulerade fallet minskar bruset i skattningarna ju större område som används.

Att skatta de lokala lutningarna över ett område innebär samtidigt att man förlorar detaljerna inom detta område. I Figur 19 visas hur felet varierar med avseende på storleken på det område över vilket lutningen skattas.; som väntat ökar detta fel ju större kärna som används då detaljer i den underliggande höjdmodellen alltmer jämnas ut.

Man kan nu börja ställa krav och dra slutsatser på datat; vill vi exempelvis att skattningen av vinkeln skall ha en felspridning mindre än 0,25 grader och att brusnivån i den underliggande höjdmodellen enligt metadata ligger på 4 cm (standardavvikelse) kräver det en kärnstorlek större än eller lika med 7, och detta innebär att avvikelserna i skattningen har en standardavvikelse på ca 0,13 grader eller mer. Utifrån detta kan man göra en initial rimlighetsanalys om höjdmodellen uppfyller kraven. Eftersom avvikelserna ligger i samma storleksordning som kravet på skattningen kan man sluta sig till att höjdmodellen ligger på gränsen till det acceptabla och att en noggrannare analys krävs. En noggrann analys kräver dock att man även undersöker hur de underliggande fördelningarna ter sig och fortplantar sig. Detta kan sällan göras på rent analytisk väg utan simuleringsbaserade metoder är oftast bättre lämpade. Detta är också knutet till vilken algoritm som används för skattningen varför det är lämpligt att kraven på att en sådan analys görs och ställs till leverantören av algoritmen. Denna analys är också nödvändig om man vill koppla spridningsmättet till sannolikheter, om bruset exempelvis är gaussiskt så motsvarar $\pm$ två standardavvikelser konfidensen 95%, dvs att sannolikheten att det sanna värdet på det man skattar faller inom givna gränser är 95%.



Figur 18. Simulerad och teoretisk fortplantning av brus i höjdmodellen uttryckt i standardavvikelse för fyra olika brusnivåer vid skattning av lutningen. Vertikal axel representerar bruset hos den skattade vinkeln, horisontell axel visar hur stort område som har använts vid skattningen.



Figur 19. Storleken på felet uttryckt i standardavvikelse som funktion av storlek på det område över vilket man skattar den lokala lutningen i x- och y-led.

Motsvarande analys som gjorts ovan bör man göra för alla ingående lager. I ett skarpt fall måste man även ta hänsyn till felaktigheter i position på lagren, både absolut men även hur väl lagren är anpassade till varandra. Man kan även använda liknande metoder för att kravställa lagren med avseende på brusnivåer, upplösningar, krav på algoritmer och liknande. I samtliga fall är metadata av yttersta vikt för att analysen skall vara möjlig att genomföra.

7 Kontakter

Under arbetets gång samverkar projektet med Geo SE på Lantmäteriet för att säkerställa relevansen hos föreslagna koncept, dela kunskaper och erfarenheter samt lägga grunden för framtida samverkan inom området.

Ett möte har också genomförts hos Saab Dynamics den 27 augusti, där en presentation av den senaste versionen av Rapid 3D Mapping gavs med lite olika exempel där systemet har använts. En genomgång gjordes också av vilka flygningar och databearbetningar Saab har genomfört det senaste året som skulle kunna vara av intresse för projektet. Diskussioner fördes också om olika kvalitetsmått samt om vilka metadata som levereras idag och vilka som skulle kunna levereras.

Kontakter har också tagits med SLB och SWECCIS för att få till ett möte för en första inblick i hur de ser på metadataproblematiken och vilka behov som kan finnas inom respektive område av metadata i allmänhet och för 3D-data i synnerhet.

Initiala kontakter har också förevarit mellan projektet och personer med anknytning till utvecklingen av geodataförsörjning i Carmenta Engine 5.

8 Förslag på fortsatt arbete

- Fortsatta diskussioner med Geo SE om metadata och innehåll i metadatafiler samt konkretisering utifrån arbetet i detta projekt. Exempel på områden som skulle kunna tas upp är:
 - FM behov/krav på metadata för olika tillämpningar
 - Militär metadataprofil
 - Specificering av de ingående metadata elementen

En vanlig kritik mot standarder går ut på att de är begränsande, så tillvida att de innebär kompromisser mellan olika intressen och att de därför inte kan täcka in "allt". Standarderna ISO 19115/19139 är dock tämligen uttrycksfulla och deras potential verkar idag utnyttjas endast blygsamt, åtminstone i Sverige. Exempelvis finns stöd för att hänvisa till publikationer som förklarar en metod som använts i dataförädlingen och möjligheter att inkludera länkar till andra dataset.

När det gäller metadata elementen så räcker det t ex inte att bara ange att data är "höjddata". Utan det måste framgå vad höjden representerar, om det t ex är markens höjd eller något annat. I aktuella standarder finns alternativet att ange "keywords", men ett bättre sätt är att göra en lista på möjliga alternativ och att definiera denna i form av en profil för höjddata.

Det använda testfallet/scenariot "HKP-landningsplats" skulle kunna utvecklas och användas som ett pilot fall.

- Hantering av heterogena dataset och den komplexitet som därav följer när det gäller att skapa innehållet i metadata för det nya datasetet är ett allt mer angeläget ämne, men bra lösningar för detta saknas. Här behöver problemen identifieras och eventuella lösningar förslås.
- Utifrån de referensdata som har byggts upp, skulle osäkerheten i ett antal vanligt förekommande COTS data vid internationella insatser kunna verifieras empiriskt.

9 Referenser

1. Tinkham, W. T., Smith, A. M. S., Hoffman, C., Hudak, A. T., Falkowski, M. J., Swanson, M. E., Gessler, P. E. "Investigating the influence of LiDAR ground surface errors on the utility of derived forest inventories", Canadian Journal of Forest Research, 42(3): 413-422, 2012
2. Vapnik, V., "The Nature of Statistical Learning Theory", Springer-Verlag, 1995

10 Förkortningar

ANSI	American National Standards Institute
CE	Circular Error
CIA	Central Intelligence Agency
CityGML	Informationsmodell för representation av gruppering av 3D urbana objekt
COTS	Commercial off-the-shelf
CSW	Catalog Service for the Web
DARPA	Defense Advanced Research Projects Agency
DEM	Digital Elevation Model
DIGEST	Digital Geographic Information Exchange Standard
DGIWG	Defence Geospatial Information Working Group
DSM	Digital Surface Model
DTM	Digital Terrain Model
EPSG	European Petroleum Survey Group
EU	Europeiska unionen
FGDC	Federal Geographic Data Committee
GEOINT	GEOspatial INTelligence
Geo SE	Försvarsmaktens stödenhet geografisk information
GIS	Geografiska Informations System
GML	Geography Markup Language
GPS	Global Positioning System
GSIP	GEOINT Structure Implementation Profile
GWG	GEOINT Standards Working Group
IMU	Inertial Measurement Unit
INSPIRE	Infrastructure for Spatial Information in the European Community
ISBN	International Standard Book Number
ISO	International Organization for Standardization
ISSN	International Standard Serial Number
LE	Linear Error
MSS	Markstridsskolan
NAP	North American Profile
NATO	North Atlantic Treaty Organization
NCGIS	National Center for Geospatial Intelligence Standards
NGA	National Geospatial-Intelligence Agency
NNH	Ny Nationell Höjdmodell

NSA	National Security Agency
NSG	National System for Geospatial-Intelligence
OGC	Open Geospatial Consortium
PDF	Portable Document Format
RH2000	Rikets Höjdsystem 2000
RMSE	Root Mean Square Error
ROE	Rules of Engagement
SAR	Synthetic Aperture Radar
SGIF	Snabb Geografisk Informationsförsörjning
SIS	Swedish Standard Institute
SLB	Stridsledningssystem Bataljon
SRTM	Shuttle Radar Topography Mission
STANAG	Standardization Agreements
SWECCIS	Swedish Command and Control Information System
SWEREF99 TM	Swedish Reference Frame 1999, Transverse Mercator
TIN	Triangulated Irregular Network
US DoD	United States Department of Defense
WCS	Web Coverage Service
WFS	Web Feature Service
WGS84	World Geodetic System 1984
WMS	Web Map Service
XML	Extensible Markup Language

Bilaga 1

Identifiera och kravställa geoinformation med tillhörande metadata

Nedan visas hur man kan arbetet utifrån systemutvecklingsmetoder för att identifiera och kravställa geoinformation och tillhörande metadata utifrån behoven hos en tillämpning.

Exemplet utgår ifrån en modifierad metod för kvalitetsdriven kravhantering, *Kvalitetsdriven kravhantering* [2]. För att ge in djupare insikt i kvalitetsdriven kravhantering förklaras de olika stegen i kvalitetsdriven kravhantering och hur dessa kommer in i den modifierade metoden.

Kvalitetsdriven kravhantering är fokuserad på identifieringen av vilka behov användare har samt hur dessa kan kopplas samman med en strukturerad kravhantering. Metoden grundar sig i ett antal aktiviteter för datainsamling med efterföljande analyser.

En grundläggande distinktion i kvalitetsdriven kravhantering är skillnaden mellan behov och krav. Behov skiljer sig från krav då de inte uttrycker vad ett system ska göra, utan vad användarna behöver stöd med. Krav å andra sidan definierar vad ett system skall och inte skall göra, även om kraven bör formuleras utan att specificera den tekniska realiseringen. Behoven är således det som ligger bakom och motiverar kraven. Vilka krav som ska ställas på systemet utgår således utifrån de behov som identifierats. Detta samband tillsammans med en strukturerad metod säkerställer en spårbarhet från de behov användare upplever och uttrycker till de krav som ska ställas på ett system.

Ett problem som kvalitetsdriven kravhantering syftar till att lösa är den informationsdiskrepans som finns vid systemutveckling mellan olika intressenter, exempelvis leverantör och användare. Leverantören har inte kunskap om vilka behov de enskilda användarna upplever och vill ha lösta. Användarna å andra sidan har inte kunskap om vad leverantören kan erbjuda och kan därmed inte specificera kraven på ett önskat system. Metoden kvalitetsdriven kravhantering innebär att användare inledningsvis i processen ses som experterna. Det är Försvarsmakten som innehar rollen som verksamhetsexpert och är källan till vilka behov som finns att uppfylla. Längre in i processen övergår expertrollen allt mer till t ex Geo SE efterhand som behoven omvandlas till krav och utformning av produkter samt vilken typ av information och metadata som krävs.

Det metodförslag som presenteras här har tagit de grundläggande begreppen och aktiviteterna från kvalitetsdriven kravhantering och applicerat dem för kravställning av metadata för tredimensionell geodata. Exempel på metadata i denna kontext kan vara information om vilket koordinatsystem som använts, vilken noggrannhet mätningarna har och tidpunkt för mätningen

De tre huvudsakliga aktiviteterna i kvalitetsdriven kravhantering är *intressentanalys*, *behovsanalys* och *kravanalys* [2]. Det övergripande resultatet av metoden är en spårbar koppling mellan slutanvändarnas behov till de faktiska kraven på geodata med tillhörande metadata.

Intressentanalys

Intressentanalysen är naturligt ett första steg och syftar till att kartlägga berörda intressenter. Intressenterna sträcker sig från slutanvändare, utvecklare till leverantören av geodata. Aktiviteten består dels av identifieringen av dessa intressenter, dels av dokumentationen av dem.

Behovsanalys

Behovsanalys syftar till att identifiera, analysera och strukturerat dokumenterar intressenters faktiska behov [2]. Dessa behov kan av intressenter vara uttalade och outtalade, samt vara tillfredsställda respektive otillfredsställda med nuvarande ansats. De tillfredsställda och outtalade behoven är mycket svåra att identifiera. En förändring kan därmed bli negativ om den leder till att outtalade och tidigare tillfredsställda behov inte längre tillfredställs.

Behovsanalys består av fyra steg: (1) insamling av data, (2) identifiering av utsagor, (3) identifiering av behov och (4) analys av behov.

Insamling av data

Det första steget syftar till att samla in information kring intressenternas situation. Detta kan ske via till exempel intervjuer, observationer, workshops och enkäter. Enligt Hallberg et al. (2006) kan det vara effektivt att helt enkelt fråga efter de problem som intressenterna upplevt. En mer formell metod med liknande ansats är *Critical Incident Technique*. Här är fokus på att be intressenterna beskriva en eller flera speciellt viktiga situationer eller händelser och hur tekniken fungerade i dessa fall. En fara med att endast fokusera på de problem som upplevs är att redan tillfredsställda behov inte uttalas. Det kan därför vara viktigt att reflektera över vilka outtalade behov som kan finnas redan vid datainsamlingen.

Uttrycker intressenterna sig i form av tekniska realiseringarna kan det vara användbart att fråga *varför* de behöver en specifik lösning. Ett flertal på varandra följande varför-frågor kan kedjas samman för att identifiera det bakomliggande behovet till att användarna önskar sig en viss lösning. Varför-tekniken är en form av utforskning av problem- och lösningsrymden och är ett sätt att djupare förstå användarens problem och behov [1]. Den lösning som användarna uttrycker behöver inte vara den som bäst uppfyller det bakomliggande behovet.

Resultatet av datainsamlingen är ostrukturerad textuell dokumentation. Ett fiktivt exempel på detta, baserat på scenariot att en underrättelseassistent i en stab har som uppgift att ta fram underlag rörande ett insatsområde ges i Tabell 1. Underlagets syfte är att identifiera lämpliga landningsplatser för helikopter.

Tabell 1 Fabricerat exempel på intervjumaterial gällande identifiering av landningsplatser

Det kritiska är ju att hitta bra plana ytor för potentiella utelandningsplatser. Det planerars före innan flygningar utifrån kartor, satellit och flyg, och observationer etc. Och där får man ju ta den informationen man har för att göra bedömningar över lämpliga platser, t.ex. underlag och lutning. Är inte platsen förberedd måste jag ju veta... är instrumenten eller information jag får något att ha? Plus minus någon grad spelar kanske ingen roll. Sen spelar ju vindförhållande etc. också roll, men talar vi om 10-20 grader felmarginal? Då går ju inte följa det.

Piloten måste ju visuellt observera platsen också men det spar ju tid att ge förslag. Speciellt om sikten är sämre, brownout eller annat eller om det är ont om tid. Om det är gott om tid kan ju mer osäkra förslag ges som kanske ligger bättre till rent infrastrukturellt. Men då måste ju piloten granska mer visuellt eller att platsen rekognoseras på plats.

Identifiering av utsagor

I detta steg är målet att identifiera intressenternas *utsagor* baserat på dokumentationen som erhöles under datainsamlingen. En utsaga är information uttryckt av en intressent som kan användas för att identifiera behov. Utsagorna ska vara koncisa och bestå av en eller ett par meningar.

Resultatet av denna analys är en strukturerad lista över utsagor, vilket exemplifieras i Tabell 2.

Tabell 2 Fabricerat exempel på identifierade utsagor

Utsaga 1	Det kritiska är ju att hitta bra plana ytor för potentiella helikopterlandningsplatser. Det planerars före innan flygningar utifrån kartor, satellit och flyg, och observationer etc. Och där får man ju ta den informationen man har för att göra bedömningar över lämpliga platser, t ex underlag och lutning.
Utsaga 2	Är instrumenten eller information jag får nåt att ha? Plus minus nån grad spelar kanske ingen roll. Sen spelar ju vindförhållande etc också roll, men talar vi om 10-20 grader felmarginal? Då går ju inte följa det.
Utsaga 3	Om det är gott om tid kan ju mer osäkra förslag ges som kanske ligger bättre till rent infrastrukturellt.

Identifiering av behov

I detta steg analyseras de identifierade utsagorna med hjälp av kundrösttabeller (eng. *Voice-of-the-Customer Table*) för att identifiera behov. Varje utsaga analyseras utifrån sex aspekter (Vem, Vad, När, Var, Varför och Hur) och slutligen identifieras vilka behov som intressenten antingen uttalat eller outtalat uttrycker i utsagan. Varje enskild utsaga kan ge upphov till flera behov.

Resultatet från analysen är en dels en lista med formulerade behov, dels kundrösttabellerna som medger en spårbarhet mellan behov och utsagor.

Tabell 3 Exempel på tre utsagor analyserad i form av kundrösttabeller

Utsaga 1	Det kritiska är ju att hitta bra plana ytor för potentiella helikopterlandningsplatser. Det planerars före innan flygningar utifrån kartor, satellit etc. Och där får man ju ta den informationen man har för att göra bedömningar över lämpliga platser, t ex underlag och lutning.
Vem	Underrättelseassistent
Vad	Identifiera lämpliga helikopterlandningsplatser
När	Inhämtning, bearbetning och sammanställning av underrättelser
Var	Stabsplats
Varför	Skapa underlag som förberedelse inför flyguppdrag
Hur	Markera ut potentiellt lämpliga helikopterlandningsplatser på karta
Behov	B0. Behov av planeringsunderlag B1. Behov av värdera olika helikopterlandningsplatser rörande underlag, marklutning och uppstickande föremål. B2. Behov av att sammanfoga olika information som kartor, satellit- och flyginformation, samt observationer.

Utsaga 2	Är instrumenten eller information jag får något att ha? Plus minus någon grad spelar kanske ingen roll. Sen spelar ju vindförhållande etc också roll, men talar vi om 10-20 grader felmarginal? Då går ju inte följa det.
Vem	Underrättelseassistent
Vad	Bedömer informationens tillförlitlighet
När	Inhämtning, bearbetning och sammanställning av underrättelser
Var	Stabsplats
Varför	För att inte skapa underlag på alltför oklara grunder. Vara förberedd på att informationen inte är 100%.
Hur	Genom att se till felmarginaler och typ av datamängd
Behov	B3. Behov av värdera tillförlitligheten hos geoinformation.

Utsaga 3	Om det är gott om tid kan ju mer osäkra förslag ges som kanske ligger bättre till rent infrastrukturellt.
Vem	Underrättelseassistent
Vad	Gör avvägningar mellan att ge förslag eller begära in mer information.
När	Inhämtning, bearbetning och sammanställning av underrättelser
Var	Stabsplats
Varför	För att inte missa lämpliga landningsplatser som ligger bättre till.
Hur	Genom att se till felmarginaler och typ av datamängd
Behov	B3. Behov av värdera tillförlitligheten hos geoinformation.

Analys av behov

I steg Analys av behov struktureras behoven efter innehåll för att hitta liknande behov. Detta kan göras med *relationsdiagram*. Behoven kan även analyseras utifrån om de är enkla eller sammansatta. Med sammansatt menas att behovet går att dela upp i flera underliggande behov. Ett sammansatt behov kan ses som en behovskategori och kan användas för att strukturera behov hierarkiskt. Genom att ensa formuleringar kan även dubletter identifieras och sorteras bort.

Genom att analysera behovens hierarki med hierarkiska diagram går det även upptäcka avsaknaden av behov. Dessa kan således kompletteras till behovsmängden.

Tabell 4 visar hur behoven har skrivits mer generellt, och istället delats in i underliggande behov. Dessa behov skulle i sin tur behöva analyseras djupare i ett verkligt fall.

Tabell 4 Exempel på identifierade behov

B0.	Behov av planeringsunderlag
B0.1	– Reglemente
B0.2	– Produktspecifikation
B0.3	– ROE
B1	Behov av värdera olika helikopterlandningsplatser enligt specifikation
B1.1	– Omvärldsmodell
B1.1.1	○ Lutning
B1.1.2	○ Markbeskaffenhet
B1.1.3	○ Uppstickande föremål
B2	Behov av att hantera olika typer av information som kartor, satellit- och flyginformation, samt observationer.
B3	Behov av att värdera tillförlitligheten hos geoinformation
B3.1	– Metadata
B3.1.1	○ Geografisknoggrannhet och -precision
B3.1.2	○ Värdenoggrannhet och -precision
B3.4	– Sensorupplösning

Kravanalys

Kravanalys syftar till att utifrån de dokumenterade behoven identifiera vad ett system skall och inte skall göra. Kraven kan delas in i funktionella och icke-funktionella krav. Ett funktionellt krav beskriver vilken funktion ett system skall utföra. Ett icke-funktionellt krav kan istället beskriva vilken information som systemet måste få tillgång till för att kunna utföra en viss funktion. Detta betyder att de icke-funktionella kraven beskriver vilken typ av information samt geoinformation med tillhörande metadata som behövs.

Kravanalys består av tre steg: (1) identifiering av krav, (2) analys av krav och (3) dokumentering av krav.

Identifiering av krav på geoinformation och metadata

Vilka krav som ska ställas på ett framtida system formuleras utifrån de identifierade behoven. Kraven formuleras så att de svarar på hur behoven kan tillfredsställas, men ska inte formuleras i form av teknisk realisering. Ett sätt att identifiera krav är att skapa *användningsfall* (eng. use cases). Användningsfall kan dels uttryckas visuellt, dels i tabellformat. Att tillsammans med intressenter skapa visuellt uttryckta användningsfall kan vara ett effektivt verktyg för att i samråd skapa en förståelse för aktuella behov och hur dessa kan uppfyllas. Användningsfallet beskriver en situation som illustrerar ett eller flera behov som en intressent har i denna specifika situation samt vad som behövs för att tillfredsställa dessa behov. Ett och samma behov kan tillfredställas på olika sätt, men kan även kräva flera krav för att tillfredsställas.

Efter hand som krav identifieras behöver en analys genomföras för att se vilken typ av geodata som behöver finnas för att uppfylla kravet. När dessa är identifierade noteras även vilken typ av metadata som är nödvändig för bearbetning av vardera typ av geodata.

I Tabell 5 visas ett exempel på ett användningsfall som beskriver en användningssituation. Denna beskrivning är baserad på identifierade behov, vilka krav detta ställer på ett system samt vilka metadata som krävs hos geodamängden.

Tabell 5 Ett förkortat exempel på hur ett användningsfall kan användas.

Namn	Framtagning av underlag för potentiella helikopterlandningsplatser baserat på geodata
Beskrivning	Operatören väljer "Identifiering av landningsplats" i sitt GIS-system. Systemet presenterar en översiktskarta för operatören. Operatören anger operationsområde. Systemet frågar en katalogtjänst efter relevant geodata över operationsområdet. Katalogtjänsten filtrerar fram data baserat på typ av data, upplösning, aktualitet och andra metadata enligt specifikation. Systemet presenterar sedan för operatören över vilka områden inom det angivna operationsområdet relevanta geodata finns, saknas samt vilka dessa är. Systemet presenterar även vilken kvalitet respektive geodata har. Operatören bedömer, utifrån bl a kvalitet och övrig metadata, vilka datalager som skall användas samt anger en inbördes prioriteringsordning mellan geodata av samma typ. Operatören anger helikoptertyp och ROE, exempelvis Hkp 13/Fred. Systemet slår upp regler för specifik helikoptertyp under angiven ROE vad gäller exempelvis maximal tillåten lutning, säkerhetsavstånd till skogsridå samt minimal tillåten markbärighet, och filtrerar geodata utifrån dessa. Systemet tar då även hänsyn till geografisk onoggrannhet. Systemet presenterar de platser som enligt datamängden troligen är lämpliga, samt vilka platser som troligen inte är lämpliga. Systemet motiverar även, utifrån regelverket, varför en given plats inte är lämplig. Systemet presenterar även för vilka områden en analys inte kan göras och vars status därmed är helt okänd. Systemet motiverar även varför en analys inte är möjlig. Operatören justerar kriterierna och ökar felmarginalerna. Systemet presenterar en ny analys över landningsplatser. Operatören markerar potentiella landningsplatser med större felmarginaler. Operatören anger att mer information krävs för dessa platser innan de brukas.
Nytta	Operatören hittar relevant information. Operatören har möjlighet att överblicka och värdera all tillgänglig relevant information samtidigt. Operatören kan beroende av behov justera vilken information som ska nyttjas och hur den presenteras. Operatören har möjlighet att ange om mer information behövs.
Behov som motiverar	B0. Behov av planeringsunderlag B1. Behov av att värdera olika helikopterlandningsplatser rörande underlag, marklutning och uppstickande föremål.

	B2. Behov av att hantera olika typer av information som kartor, satellit- och flyginformation, samt observationer.
	B3. Behov av att värdera tillförlitligheten hos geodata
Primär aktör	Operatör (Underrättelseassistent)
Funktionella krav	Systemet skall presentera en översiktskarta. Systemet skall ge Operatören möjlighet att ange operationsområde. Systemet skall kunna avgöra vilka geodata som är relevanta och som uppfyller specifikationerna enligt regelverk för helikopterlandningsplatser. Systemet skall kunna kommunicera med en katalogtjänst. Systemet skall kunna efterfråga och hämta geodata från geodatatjänster. Systemet skall presentera över vilka områden inom det angivna operationsområdet relevanta geodata finns. Systemet skall presentera över vilka områden inom det angivna operationsområdet relevanta geodata saknas. Systemet skall presentera vilken kvalitet respektive geodata har. Systemet skall ge Operatören stöd i att välja vilka datalager som skall användas samt ange en prioritering. Systemet skall ge Operatören stöd att granska metadata. Systemet skall ge Operatören stöd i att ange helikoptertyp och ROE. Systemet skall avgöra vilka regler som gäller för specifik helikoptertyp under angiven ROE och filtrerar geodata utifrån dessa. Systemet skall ta hänsyn till geografisk onoggrannhet och presentera denna för Operatören. Systemet skall presentera de platser som utifrån helikoptertyp och ROE är lämpliga. Systemet skall presentera de platser som utifrån helikoptertyp och ROE inte är lämpliga. Systemet skall motivera varför en given plats inte är lämplig. Systemet skall presentera över vilka områden en analys inte kan göras och även varför en analys inte är möjlig. Systemet skall ge Operatören stöd i att justera kriterierna och öka felmarginalerna och därefter presentera en ny analys över potentiella landningsplatser. Systemet skall ge Operatören möjlighet att markera potentiella landningsplatser vilka uppfyller specifikationerna. Systemet skall ge Operatören möjlighet att markera potentiella landningsplatser vilka inte uppfyller specifikationerna. Systemet skall kunna sammanställa information om valda helikopterlandningsplatser. Systemet skall även ge Operatören stöd i att annotera valda

helikopterlandningsplatser med text.

Systemet skall hantera olika typer av filter, t.ex. för att extrahera lutning ut höjdinformation eller skattadiametern på homogent område.

Icke funktionella krav

Information	Planeringsunderlag (operationsområde, helikoptertyp(Maximal tillåten lutning, Minimal markbärighet, Dimensioner), ROE(Avstånd till uppstickande föremål, Säkerhetsmarginaler på lutning och markbärighet))
	Regelverk för helikopter under angiven ROE
Geodata	Utiifrån Produktspecifikation helikopterlandningsplats data som kan ge markbärighet, marklutning och uppstickande föremål.
Metadata som krävs för geodata	Geografisk projektion Geografisknoggrannhet och –precision Värdenoggrannhet och –precision Tidstämpel Geografisk utbredning Projektion Sensor/process historik Särskilda egenskaper för data Typ av data Verklig upplösning på data

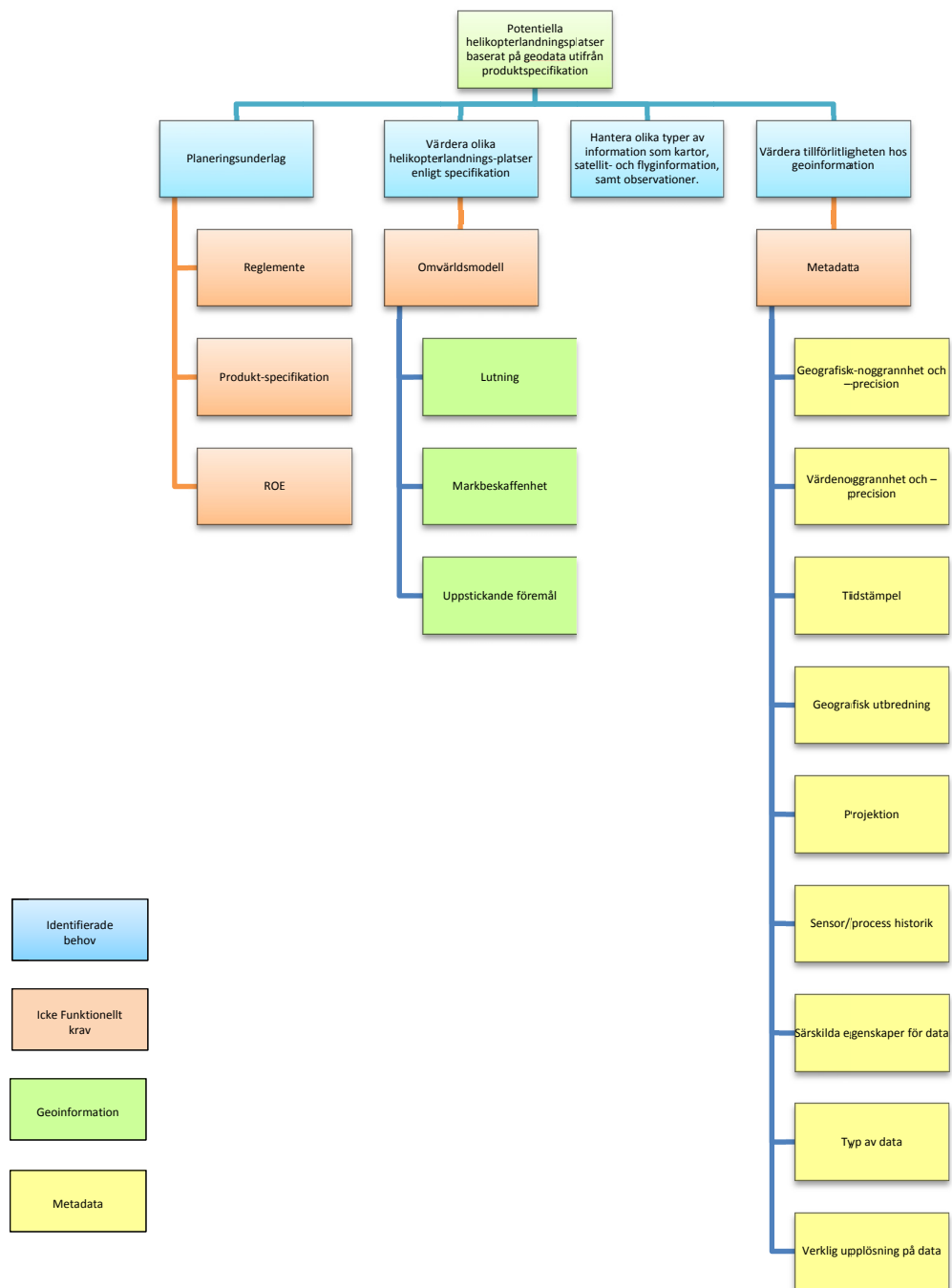
Analys av information

Under analysen av de icke-funktionella kraven görs en liknande analys som för behoven, alltså att kraven struktureras för att identifiera grupperingen, likartade krav eller dubletter.

De strukturerade icke-funktionella kraven kan återges i en trädstruktur med kopplingar till de tidigare identifierade behoven och det övergripande målet (Figur 1).

Dokumentering av information

Det är viktigt att dokumenteringen av det framkomna informationsbehovet görs på ett ändamålsenligt sätt så att det blir ett stöd i framtagande av vilka typer av produkter och tjänster som borde tillhandhållas och vilka geodata och metadata som behövs för att understödja dessa. Dokumentationen måste också bibehålla spårbarheten tillbaka till behoven och det övergripande målet. Detta för att underlätta hanteringen av framtida förändringar i kedjan - behov till geodata - och hur detta påverkar produkter och tjänster.



Figur 1 Exempel på icke funktionella krav.

Sammanfattning

Exemplet visar hur systemutvecklingsmetoder kan användas för att kravställa geoinformation samt tillhörande metadata. Exemplet utgår ifrån en modifiering av metoden *Kvalitetsdriven kravhantering* [2] som används vid systemutveckling för att bli lösa den informationsdiskrepans som finns mellan olika intressenter, exempelvis leverantör och användare.

De tre huvudsakliga aktiviteterna i metoden är intressentanalys, behovsanalys och kravanalys. Metoden identifiera vilka behov användare har genom en föreskriven behovsanalys, vilket sedan kopplas samman med kravhantering. I systemutveckling används kravanalys primärt för att identifiera vilka krav som ska specificera ett tekniskt system. I det här fallet är det primära syftet inte att specificera ett tekniskt system, utan identifiera vilken geoinformation med tillhörande metadata som behövas i ett geodatasystem.

Genom den modifierade metoden visas att en spårbar koppling mellan slutanvändarnas behov till de faktiska kraven på geodata med tillhörande metadata kan uppnås via en systemutvecklingsmetod.

Referenser

1. Löwgren, J. & Stolterman, E. (2004). "Design av informationsteknik: Materialet utan egenskaper" (2:a uppl.). Studentlitteratur, Lund.
2. Hallberg, N., Ölvander, C. & Törne, A. (2006). "Behovs- och kravanalys avseende tekniskt beslutsstöd för operationer i urban terräng". FOI-R--2037--SE. Totalförsvarets forskningsinstitut, Linköping.