

Analys av den norska standarden  
**Kontroll av geodata**

Patric Jansson & Clas-Göran Persson

Teknisk rapport 2013:1





## Förord

Det här är den andra, officiella versionen av denna rapport. I förhållande till Version 1 har ett antal justeringar och tillägg gjorts, baserade på en "kvalitetssäkring" i form av en studieresa till Kartverket i Norge den 10-11 september.

Rapporten utgör den första i den nya skriftserien "Tekniska Rapporter" inom ramen för HMK (Handbok i mät- och kartfrågor). Där kommer analyser, bakgrundsfakta, referensmaterial etc. att publiceras, vilket förhoppningsvis ökar förståelsen och gör det möjligt att gå mer "rakt på sak" i de regelrätta handbokstexterna. [Samlade förord](#)

## Abstract

This report describes the Norwegian standard "Kontroll av geodata" (Control of Geodata). The standard is presented and its usefulness for the Swedish ongoing work with HMK (Handbok i mät- och kartfrågor; Handbook in Surveying and Mapping) is analysed.

A preliminary conclusion from the authors is that the standard could be implemented to Swedish conditions, without the need for any larger adjustments.

**Key words:** geodata, quality control, standardisation, tolerances.

## Författarnas kontaktuppgifter

Patric Jansson

Kungliga Tekniska Högskolan (KTH)  
Drottning Kristinas väg 30  
SE - 100 44 Stockholm  
[patric.jansson@abe.kth.se](mailto:patric.jansson@abe.kth.se)  
+46-8-790 7376

Clas-Göran Persson

Kungliga Tekniska Högskolan (KTH)  
Drottning Kristinas väg 30  
SE - 100 44 Stockholm och  
Lantmäteriet  
SE - 801 82 Gävle  
[clas-goran.persson@lm.se](mailto:clas-goran.persson@lm.se)  
+46-70-557 6037

# Innehållsförteckning

<b>1</b>	<b>Syfte</b> .....	<b>5</b>
<b>2</b>	<b>Det norska standardiseringsarbetet</b> .....	<b>5</b>
<b>3</b>	<b>Presentation och analys av dokumentet "Kontroll av geodata"</b> .....	<b>5</b>
3.1	Syftet med dokumentet (version 2007).....	5
3.2	Innehåll.....	6
3.3	Kontrollprocessen.....	7
3.4	Kontrollområden.....	8
3.5	Stickprov och stickprovsstorlek.....	9
3.6	Kontroll av grova fel, fullständighet och objektklassificering .....	10
3.7	Kontroll av lägesosäkerhet .....	12
3.8	Kontroll av systematik .....	12
<b>4</b>	<b>Användbarhet i HMK-arbetet</b> .....	<b>14</b>
4.1	Möjligheter till harmonisering .....	14
4.2	Användning av termen "tolerans" .....	15
<b>5</b>	<b>Tillämpningsexempel</b> .....	<b>16</b>
5.1	Produktspecifikationens krav .....	16
5.2	Test av objektklassificering.....	17
5.3	Test av fullständighet .....	17
5.4	Test av lägesosäkerhet, systematik och grova fel.....	17
	Grova fel .....	17
	Lägesosäkerhet .....	17
	Systematik .....	18
<b>6</b>	<b>Epilog</b> .....	<b>18</b>
6.1	Reflektioner från Norgebesöket.....	18
6.2	Slutord .....	20
<b>7</b>	<b>Referenser</b> .....	<b>20</b>
	<b>Bilaga 1: Exempel på kvalitetskrav vid kontrollmätning</b> .....	<b>21</b>
	<b>Bilaga 2: Test av grova fel, mätosäkerhet och systematik i 1 D, 2D och 3D</b> .....	<b>22</b>
	<b>Bilaga 3: Förslag till harmonisering av terminologin beträffande kvalitetskrav</b> .....	<b>26</b>

## 1 Syfte

Denna rapport innehåller beskrivning och analys av den norska standarden "Kontroll av geodata". Syftet är att bedöma dess användbarhet i motsvarande svenska arbete inom ramen för HMK (Handbok i mät- och kartfrågor). Arbetet har initierats av Lantmäteriet.

## 2 Det norska standardiseringsarbetet

Norge har under lång tid satsat på samordning och standardisering inom geodataområdet.

- **SOSI** (SOSI-standard): "Samordnet Opplegg for Stedfestet Informasjon" är ett system för standardiserad beskrivning av digitala geodata.
- **Geovekst** är ett geodatasamarbete som syftar till att geografisk information samlas in en gång och ajourhålls av en organisation. Samarbetsavtalet ingicks 1992 mellan Statens vegvesen, Energiforsyningens Fellesorganisasjon, Kommunenes Sentralforbund, Kartverket, Telenor och Landbruks- og matdepartementet.
- **Norge Digitalt** är ett brett samarbete mellan verksamheter som har ansvar för att skaffa fram lägesbunden information och/eller som är stora användare av sådan information. Norge digitalt-samarbetet har sin förankring i Stortingsmelding nr. 30 (2002-2003): "Norge digitalt - et felles fundament for verdiskaping".
- **FKB** (Felles KartdataBase) består av strukturerade vektor-data. Data är specificerade i FKB-standarderna FKB-A, FKB-B, FKB-C och FKB-D, som ska täcka behovet av gemensamma kartdatabaser i de olika områdestyper som är definierade i Geodatastandard (se nästa avsnitt).

## 3 Presentation och analys av dokumentet "Kontroll av geodata"

### 3.1 Syftet med dokumentet (version 2007)

Standarden "Kontroll av geodata" är koordinerad med "Kart og geodata", och båda dokumenten bygger på "Kvalitetssikring av oppmåling, kartlegging og geodata" (Geodatastandard).

Kvalitetskraven som resultaten av kontrollerna ska jämföras med finns i olika *produktspecifikationer*, t.ex. "SOSI Del 3, Produktspesifikasjon for Felles KartdataBase (FKB)". "Kontroll av geodata" hade varit i bruk i 5 år när den reviderades 2007.

Syftet är att bidra till hög kvalitet på geodata. Standarden specificerar hur kontroll av geodata ska utföras för att vara statistiskt trovärdig. Tanken är att datamängder som håller specificerad kvalitet med stor sannolikhet kommer att klara standardens krav vid kontroll. Begreppet "produkt" används såväl för dataleveranser från utförare till beställare som för databaser och uttag ur sådana till en användare.

Kontrollerna kan svara på frågor om huruvida kvalitetskraven enligt specifikationen är uppfyllda, t.ex. vad gäller aktualitet, fullständighet m.m. Men de ger inte svar på frågor av typen "Kan datamängden användas som underlag för en detaljplan?"

Viss kunskap i det norska språket förutsätts i det följande.

## 3.2 Innehåll

Dokumentet har följande innehåll:

- **Formål og omfang.** Förhållande till geodatastandard, produktspecifikation och ansvarsfördelning beträffande kontroller.
- **Referanser, definisjoner og forkortelser.** Några termer: målning (mätning), fasitmålning (kontrollmätning med noggrannare metod), synfaring (fältkontroll), avvik/standardavvik (i stället för "fel" och "medelfel")
- **Kontrollprocessen.** Kvalitetsparametrar, kvalitetsmått, omfång, val av kontrollmetod samt genomförande och rapportering av kontroll.
- **Omfang av kontrollen.** Kontrollområden och urval av objekttyper. Fullständig kontroll och stickprovskontroll.
- **Klassifikasjon av kontrollmetoder.** Direkta och indirekta kontrollmetoder.
- **Målinger og beregninger.** Olika typer av kontroller: lägesosäkerhet, kvantitativ och kvalitativ attributosäkerhet samt fullständighet.
- **Sammenlikning av målt kvalitet mot krav.** Signifikans, hypotestest, jämförelse med toleranser samt godkännande-/avvikelsehantering.

- **Rapportering.** Kontrollrapporter och rapportering i meta-data.

med bilagorna

- **Tillegg A** (normativt): Oversikt over kvalitetselement og kvalitetsmål.
- **Tillegg B** (informativt): Aktuelle metoder for kontroll av FKB-data.
- **Tillegg C** (informativt): Utkast til kontrollplaner og rapport-skjemaer for kartleggingsprosjekt.
- **Tillegg D** (informativt): Eksempel på kontrollplaner og standardkontroll av fotogrammetrisk produserte FKB-data

### 3.3 Kontrollprosessen

Kontrollprosessen består av momenten "Upprätta kontrollplan" (Lage kontrollplan) samt "Genomförande av kontroll" (Gjennomføring av kontroll), se Tabell 1.

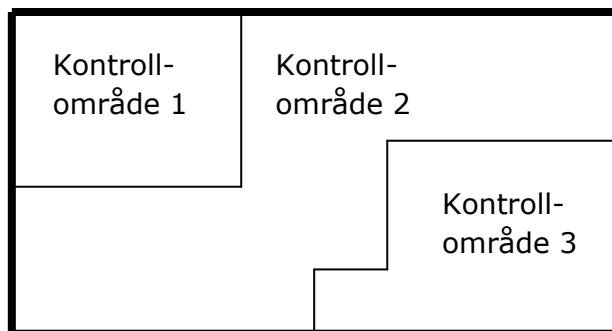
**Tabell 1.** Kontrollprocessens olika steg. Numrering enligt standarden "Kontroll av geodata".

3.1 Lage kontrollplan	
3.1.1 Velg datasett	Eksempel: Kontroll skal utføres av FKB-B kartlegging. Det velges ut to datasett for kontroll: "FKB Bygning" og "FKB LedningVa".
3.1.2 Velg kvalitetselement	Eksempel: Kontroll av fullstendighet til objekttypene ElvBekk (bekker) og Innsjø i FKBdatasettet "FKB Vann".
3.1.3 Velg kvalitetsmål	Eksempel: Andel manglende forekomster av objekttypen Bygning, oppgitt i %.
3.1.4 Identifiser toleranser	Eksempel: Krav til fullstendighet for objekttypen Bygning konstruert fra 1:8000-bilder er oppgitt som toleranse for manglende bygninger lik 0,5 % (FKB-B og fullstendighetsklasse 1).
3.1.5 Velg omfanget av kontrollen	Eksempel: Kontrollere objekttype Mønelinje innenfor 2 stk. 1:1000 kartblad, totalt 125 bygninger.
3.1.6 Velg kontrollmetode	Eksempel: Velg kontrollmetoden synfaring.

3.2 Gjennomføring av kontroll	
3.2.1 Utfør kontrollmålinger og beregn verdier for kvalitetsmål	Eksempel: En stikkprøve inneholder 125 bygninger. Fullstendighet av objekttypen Bygning skal kontrolleres. Det utføres synfaring som viser at datasettet mangler 2 bygninger som fins i terrenget (fasiten), altså mangler 2 bygninger av totalt 127. Prosentandel manglende bygninger beregnes til å være $2/(125+2) \cdot 100 \% = 1,6 \%$ .
3.2.2 Sammenlikne beregnede verdier for kvalitetsmål med toleranser	Eksempel: Andel manglende bygninger: 1,6 %. Toleranse: 0,5 %. Man må utføre test for konklusjon.
3.2.3 Godkjenning og grovfeil-/avviksbehandling	Godkjenning og grovfeil-/avviksbehandling skal skje i henhold til avsnitt 7.4.
3.2.4 Rapportering	Kontrollen rapporteres, se kapittel 8. Tillegg C, D og F gir eksempler på rapporter.

### 3.4 Kontrollområden

Datamængden delas geografisk in i *kontrollområden* på ett sådant sätt att de tillsammans (vanligtvis) täcker hela kartläggningsområdet, se Figur 1. Om t.ex. kontrollområde 1 kontrolleras gäller utfallet för hela kontrollområde 1, men bara för detta. Dvs. om kontrollen överskrider givna toleranser i ett stickprov så underkänns hela datamængden i det berörda kontrollområdet. Hur en sådan brist åtgärdas är en sak mellan beställare och utförare att avtala om. Minst två objekttyper ska kontrolleras.



**Figur 1.** En indelning i tre kontrollområden som tillsammans täcker kartläggningsområdet (fritt efter "Kontroll av geodata").



### 3.5 Stickprov och stickprovsstorlek

Vanligen är det orimligt att kontrollera samtliga objekt inom ett kontrollområde. I de fall så sker är det normalt en maskinell-/automatisk kontroll, t.ex. av topologiska egenskaper (slutna nätverk m.m.) och att attributvärden ligger inom tillåtna gränser. Totalkontroll tillämpas dock även när det totala antalet objekt är litet.

I dokumentet finns angivet hur stora stickprov som (minst) bör tas, se Tabell 2. Stickprovsstorleken beror på hur många objekt av den aktuella objekttypen som finns i datamängden, men till del också på vad som ska kontrolleras. Där företeelsen är *kvalitativ*, t.ex. test av grova fel eller fullständighet (finns/finns inte), krävs större stickprov än vid t.ex. kontroll av standardavvikelse eller systematik. I det senare fallet bedöms man få mer information eftersom avvikelserna ligger på en *kvantitativ* intervallskala, dvs. man kan "steglöst" mäta om avvikelsen är stor eller liten.

Vi har inte hittat något matematiskt uttryck för förhållandet mellan antal objekt och stickprovsstorleken. I "Kontroll av geodata" hänvisas till ISO-standarderna 2859 och 3915.

**Tabell 2.** Minsta stickprovsstorlek (ur "Kontroll av geodata").

Antal förekomster av objekttypen i kontrollområdet		Stickprovsstorlek (kvalitativ kontroll)	Stickprovsstorlek (kvantitativ kontroll)
Från	Till		
1	8 (5)	Alla objekt	Alla objekt
9 (6)	50	8	5
51	90	13	7
91	150	20	10
151	280	32	15
281	400	50	20
401	500	60	25
501	1200	80	35
1201	3200	125	50
3201	10 000	200	75
10 001	35 000	315	100
35 001	150 000	500	150
150 001	500 000	800	200
> 500 000		1250	200

Antalet objekttyper som måste kontrolleras kan bestämmas på följande sätt (exempel från "Kontroll av geodata"):

- Man har valt att kontrollera "FKB Väg" i ett kontrollområde. Det finns totalt 757 016 punkter i detta dataset. Det rör sig om en kvalitativ kontroll, varför Tabell 2, kolumn 3 ger den totala stickprovsstorleken  $n = 1250$  punkter/objekt.
- Man börjar t.ex. med objekttypen Dekkekant som visar sig innehålla 236 841 punkter, dvs. stickprovet ska omfatta  $n_1 = 800$  punkter. Det återstår alltså  $1250 - 800 = 450$  punkter.
- Därefter väljer man objekttypen Ferist, som det bara finns 20 stycken av inom kontrollområdet. Det ger stickprovet  $n_2 = 8$ , och det återstår 442 punkter för att komma upp i det stipulerade värdet 1250.
- Man väljer så objekttypen GangSykkelveg, Den består av 69 611 punkter som ger  $n_3 = 500$ . Totalt får vi alltså  $n_1 + n_2 + n_3 = 800 + 8 + 500 = 1308$ , vilket uppfyller kraven.
- Om någon av objekttyperna har för dålig kvalitet underkänns hela FKB-datasettet Väg.

### 3.6 Kontroll av grova fel, fullständighet och objektklassificering

Tabell 3 gäller för kontroll av grova fel och fullständighet/objekt-klassificering ur stickprov. Kontrollen går helt enkelt ut på att jämföra antalet hittade fel med tabellvärdena för antalet tillåtna fel. De gulmarkerade siffrorna i tabellen är subjektivt satta medan övriga erhålls ur formeln för *binomialfördelningens kumulativa fördelningsfunktion*

$$F(k) = P(X \leq k) = \sum_{i=0}^k \binom{n}{i} \cdot (1-p_o)^i (p_o)^{n-i} \quad (1)$$

där  $P(\square)$  betecknar sannolikhet och  $X$  betecknar ett "lyckat utfall".  $p_o$  är den andel felaktigheter som accepteras i produktspecifikationen, dvs.  $P(X) = 1 - p_o$

Först bestäms storheten  $k_o$  som det minsta heltal (en integer  $\leq n$ ) för vilket

$$F(k_o) = P(X \leq k_o) \geq 5\% \quad (2)$$

dvs.

$$F(k_o - 1) = P(X \leq k_o - 1) < 5\% \quad (3)$$

**Tabell 3.** Kontroll av grova fel och fullständighet/objektklassificering ur stickprov. Antalet grova fel eller fel redovisade objekt ska vara mindre än tabellvärdena. Vid kontroll av fullständighet inkluderas såväl övertäckning (för många objekt) som undertäckning (för få objekt) – allt enligt produktspecifikationen. (Ur "Kontroll av geodata".)

Antal förekomster av objekttypen i kontrollområdet		Stickprovstorlek ( $n$ )	$p_o$ = andel tillåtna fel enligt produktspecifikationen					
			0.5%	1.0%	2.0%	3.0%	4.0%	5.0%
Från	Till		Förkastningsgräns					
1	8	Alla objekt	1	1	1	1	1	1
9	50	8	1	1	1	2	2	2
51	90	13	1	1	2	2	2	3
91	150	20	1	2	2	3	3	4
151	280	32	1	2	3	3	4	4
281	400	50	2	3	3	4	5	6
401	500	60	2	3	4	5	6	7
501	1200	80	3	3	5	6	7	8
1201	3200	125	3	4	6	8	10	11
3201	10 000	200	4	6	8	11	14	16
10 001	35 000	315	5	7	12	16	20	23
35 001	150 000	500	6	10	16	23	28	34
150 001	500 000	800	9	14	24	33	42	51
> 500 000		1250	12	20	34	49	63	76

Det innebär att sannolikheten för att få  $n - (k_o - 1)$  stycken grova fel är mindre än 5 %. Dvs. vår tolerans blir

$$\text{tolerans} = n - (k_o - 1) = n - k_o + 1 \quad (4)$$

I programmeringsspråket MATLAB skrivs detta

$$\text{tolerans} = n - \text{binoinv}(0.05, n, 1 - p_o) + 1 \quad (5)$$

där *binoinv* står för "Binomial inverse cumulative distribution function". Ett alternativ till Tabell 3 är alltså att producera gränsvärdena m.h.a. ett enkelt datorprogram. Eftersom man har valt att antalet grova fel ska vara mindre än tabellvärdena blir i realiteten kraven strängare än binomialfördelningens teoretiska värden för angivna konfidensgrader.

### 3.7 Kontroll av lägesosäkerhet

Tabell 4 gäller för kontroll av standardosäkerhet (standardavvikelse) vid kontroll av produktspecifikationens kvalitetskrav.

Högra kolumnen kan tillräckligt bra approximeras med  $\sqrt{F_{0.05,n-1,\infty}} \approx (0,96 + (n-1)^{-0.4})$  och testet lyder

$$s \leq \sigma \cdot \sqrt{F_{0.05,n-1,\infty}} \approx u(l) \cdot (0,96 + (n-1)^{-0.4}) \quad (6)$$

där  $s$  är stickprovets standardavvikelse och  $u(l)$  - eller  $\sigma$  - är den standardosäkerhet som stipuleras för den aktuella objekttypen enligt produktspecifikationen. Detta är exakt samma kontroll som i HMK-Stommätning (Lantmäteriverket, 1996) Tabell A.7 och A.12! Se även HMK-Introduktion, Bilaga A.

**Tabell 4.** Kontroll av standardavvikelse. Totala antalet förekomster av den objekttyp som ska kontrolleras bör vara > 25 st. (Ur "Kontroll av geodata".)

Antal förekomster av objekttypen i kontrollområdet		Stickprovstorlek ( $n$ )	$\sqrt{F_{0.05,n-1,\infty}}$
Från	Till		
26	50	5	1.54
51	90	7	1.45
91	150	10	1.37
151	280	15	1.30
281	400	20	1.26
401	500	25	1.23
501	1200	35	1.20
1201	3200	50	1.16
3201	10 000	75	1.13
10 001	35 000	100	1.12
35 001	150 000	150	1.09
150 001	500 000	200	1.08
> 500 000		200	1.08

### 3.8 Kontroll av systematik

Tabell 5 gäller för kontroll av systematik, t.ex. systematisk avvikelse i höjd beräknad som medelavvikelse för stickprovet.

**Tabell 5.** Kontroll av systematisk avvikelse. Antalet förekomster av respektive objekttyp bör inte vara mindre än 26. (Ur "Kontroll av geodata".)

Antal förekomster av objekttypen i kontrollområdet		Stickprovsstorlek (n)	$t_{0.95,n-1}$	$\frac{t}{\sqrt{n}}$
Från	Till			
26	50	5	2.13	0.95
51	90	7	1.94	0.73
91	150	10	1.83	0.58
151	280	15	1.76	0.45
281	400	20	1.73	0.39
401	500	25	1.71	0.34
501	1200	35	1.69	0.29
1201	3200	50	1.68	0.24
3201	10 000	75	1.67	0.19
10 001	35 000	100	1.66	0.17
35 001	150 000	150	1.66	0.14
150 001	500 000	200	1.65	0.12
> 500 000		200	1.65	0.12

Kontrollen går ut på att jämföra stickprovets beräknade medelavvikelse  $\bar{x}$  med den högsta systematiska avvikelse  $\mu$  ( $\geq 0$ ) som produktspecifikationen tolererar. Detta test formuleras i "Kontroll av geodata" som

$$\bar{x} - s \cdot \frac{t}{\sqrt{n}} \leq \mu \quad (7)$$

där  $s$  är stickprovets skattade standardosäkerhet, eller standardavvikelse, och  $\frac{t}{\sqrt{n}}$  ges av Tabell 5. (Detta test ifrågasätts i Bilaga 2 – men förutom det som framförs där så måste det vara absolutvärdet av medeltalet som ska ingå in Formel 7, dvs.  $|\bar{x}|$  i stället för  $\bar{x}$ .)

Hur man gör vid 2D-kontroller, t.ex. av systematiska avvikelser i plan, är något oklart. Vi förutsätter dock att man även då tillämpar Formel 7 med den radiella avvikelsen i stället för  $|\bar{x}|$  och standardosäkerheten i plan  $s_{plan} = \sqrt{s_{Northing}^2 + s_{Easting}^2}$  i stället för  $s$ , även om detta inte är helt korrekt ur statistisk synvinkel.

## 4 Användbarhet i HMK-arbetet

Det finns flera faktorer som talar för att anamma det norska konceptet i HMK-arbetet:

- Det är beprövat och bygger på internationella ISO-standarder, samt på vedertagna statistiska metoder och tankegångarna inom GUM.
- Det harmonierar med det traditionella HMK-konceptet, men innehåller användbara utvidgningar.
- Det finns fördelar med att ha en gemensam syn på geodatakvalitet inom Norden, som ju är en och samma marknad för producenterna; det blir enklare och därmed billigare för alla.
- Vi tjänar tid.
- Det finns egentligen bara en hake: terminologin. Och då avses inte språkliga skillnader mellan norska och svenska utan, främst, den olikhet som finns i användningen av termen "tolerans", som finns i båda språken.

### 4.1 Möjligheter till harmonisering

Grunderna för det norska konceptet är:

1. Det är stickprov ur datamängden som analyseras vid kvalitetskontrollen. Det finns regler för stickprovets storlek i förhållande till datamängdens totala storlek, se Tabell 2-5.
2. Mätosäkerhetskraven i produktspecifikationen utgår från olika objekts standardosäkerhet, dvs.  $\sigma$  eller med svenska GUM-beteckningar  $u(\cdot)$ . Kontrollen av att detta är uppfyllt görs via F-testet i Tabell 4; konfidensnivå 95 %.
3. Kontrollmetoden bör ha en standardosäkerhet som är högst  $\frac{1}{3}\sigma$ . Då kan den betraktas som felfri, i annat fall måste hänsyn tas till mätosäkerheten vid kontrollen.
4. Grova fel är sådana som är större än  $3\sigma$ . Det får förekomma ett antal sådana i ett stickprov beroende på stickprovets storlek, se Tabell 3. Denna tabell används även för analys av fullständigheten, och i båda fall beror gränsvärdena på vilka krav på acceptabel felfrekvens som produktspecifikationen anger.
5. Test av systematisk avvikelse görs med ett t-test enligt Tabell 5; avsedd konfidensnivå 95 %.

Om detta kan anföras:

- Punkt 1 skulle bli ett tillägg till HMK-konceptet som inte alls påverkar övriga delar. Där finns hittills inga principer angivna för ett stickprovs storlek. Innan vi anammar denna princip bör dock utredas hur stort genomslag riktlinjerna har haft i praktiken. Stickproven ter sig rätt så stora och frågan är i vilken grad de har tillämpats.
- Punkt 2 är identisk med resonemanget inom HMK.
- Hur mätosäkerheten i kontrollmetoden påverkar de olika testerna utreds i HMK-Introduktion, som dock inte explicit kräver en standardosäkerhet på max  $\frac{1}{3}\sigma$ .
- Definitionen av grovt fel i punkt 4 är lika. Men HMK tillåter inga sådana alls medan "Test av geodata" gör det i viss grad; dock ska de fel som hittas korrigeras. En rimlig modifiering.
- Att använda Tabell 2 även för kontroll av fullständighet skulle bli ett nytillägg i HMK, men ett odramatiskt och välkommet sådant.
- Ett test av systematik finns även i HMK-Introduktion, Bilaga A, men den görs litet annorlunda. Det finns dock anledning att överväga en harmonisering med "Kontroll av geodata".
- Den underliggande konfidensnivån i "Test av geodata" är 95 %, precis som i HMK. Men i Norge tillämpas inte  $2\sigma$  som *varningsgräns* på samma sätt som i HMK.
- Man har gjort en viss GUM-anpassning av terminologin men termen *tillförlitlighet* hanteras helt annorlunda än i det svenska, Baarda-influerade konceptet. Det är dock av mindre betydelse eftersom den snarlika termen *kontrollerbarhet* och *k-tal* har fått större genomslag i Sverige och är därför central.
- "Kontroll av geodata" tillämpar inte det klassiska HMK-konceptet med tre nivåer I( $1\sigma$ ), II( $2\sigma$ ) och III ( $3\sigma$ ), se HMK-Introduktion sid. 23-24. Samtliga nivåer finns dock med indirekt.

## 4.2 Användning av termen "tolerans"

I "Kontroll av geodata" benämns de standardosäkerheter som stipuleras i produktspecifikationen för *toleranser*. Sedan gör man ett statistiskt test av om dessa toleranser är uppfyllda genom att jämföra med ett *gränsvärde*, se t.ex. Tabell 4 i avsnitt 3.7. I HMK har språkbruket varit att gränsvärdena är det som benämns toleranser

(tidigare felgränser). Viss avvikelse finns i HMK-Detaljmätningsterminologi, men hur som helst blir det litet luddigt och risken för missförstånd är uppenbar.

Ett första försök till förtydligande av detta presenteras i Bilaga 3. De kvalitetskrav som anges i produktspecifikationen benämns där *specificerad kvalitet* och den som erhålls vid kontrollmätningen för *erhållen kvalitet*. Den erhållna kvalitets tillåtna avvikelse från den specificerade benämns *tolerans*.

## 5 Tillämpningsexempel

Vi väljer att presentera förslaget till tillämpning i HMK med ett fiktivt tillämpningsexempel. Vi anammar "Kontroll av geodata" rakt av, trots de invändningar som redovisas i Bilaga 2; dessa synpunkter bör diskuteras vidare innan några förändringar görs.

### 5.1 Produktspecifikationens krav

Tänkt produktspecifikation – eller egentligen *dataproduktspecifikation* enligt nutida terminologi – har följande specificerade kvalitet:

- **Felklassificerade objekt**, byggnader:  $p_o = 0.03$  (3 %)
  - **Fullständighet**, stödmurar:  $p_o = 0.05$  (5 %)
  - **Lägesosäkerhet** (i plan), gränspunkter:  $\sigma = 20$  mm (standardosäkerhet,  $u(\cdot)$ )
- Ingen **systematisk avvikelse**:  $\mu = 0$
- Grova fel**:  $p_o = 0.01$  (1 %)

Tre kontrollområden, som tillsammans täcker datamängdens geografiska utbredning, läggs ut. Tre olika exempel på stickprov redovisas i Tabell 6.

**Tabell 6.** Stickprovsstorlekar för tillämpningsexemplet. Vad gäller kontrollen av gränspunkter bör dock stickprovet innehålla 20 objekt eftersom kontrollen av lägesosäkerhet, systematik och grova fel är samordnad.

Objekttyp	Typ av kontroll	Antal förekomster i kontrollområdet	Stickprovsstorlek ( $n$ )
Byggnader	Objektklassificering (kvalitativ)	856	80
Stödmurar	Fullständighet (kvalitativ)	58	13
Gränspunkter	Lägesosäkerhet och systematik (kvantitativ) grova fel (kvalitativ)	132	10 (20) 20



## 5.2 Test av objektklassificering

För ett stickprov på 80 objekt och med den specificerade kvalitén  $p_o = 0.03$  (3 %) blir toleransen enligt Tabell 3

**Tolerans, objektklassificering: antalet felklassificeringar < 6 st.**

Vid kontrollen fann man 5 Uthus som hade klassificerats som Bostadshus. Eftersom  $5 < 6$  så godkänns det aktuella kontrollområdet i detta avseende. Felklassificeringarna ska dock rättas till.

## 5.3 Test av fullständighet

För ett stickprov på 13 objekt och med den specificerade kvalitén  $p_o = 0.05$  (5 %) blir toleransen enligt Tabell 3

**Tolerans, fullständighet: antalet brister < 3 st.**

Vid kontrollen saknades 3 stödmurar som borde ha varit med enligt produktspecifikationen. Eftersom  $3 \geq 3$  så underkänns hela det aktuella kontrollområdet i detta avseende. Komplettering med de saknade stödmurarna ska också ske.

## 5.4 Test av lägesosäkerhet, systematik och grova fel

Testen av lägesosäkerhet, systematik och grova fel samordnas. Samtliga utgår från produktspecifikationens specificerade kvalitet för standardosäkerheten i plan<sup>1</sup> ( $\sigma$ ). Detektering av ev. grova fel bör ske först.

### Grova fel

Avvikelser större än  $3\sigma$  definieras som grova fel. Det ger i vårt fall gränsvärdet 60 mm för radiella fel. För ett stickprov på 20 objekt och med den specificerade kvalitén  $p_o = 0.01$  (1 %) blir toleransen enligt Tabell 3

**Tolerans, grova fel: antalet avvikelser större än  $3\sigma < 2$  st.**

Vid kontrollen hittades ett grovt fel. Eftersom  $1 < 2$  så godkänns hela det aktuella kontrollområdet i detta avseende. Det grova felet ska dock rättas till.

### Lägesosäkerhet

För ett stickprov på 20 objekt, med den specificerade kvalitén att standardosäkerheten i plan ska vara max 20 mm, blir toleransen enligt Tabell 4

---

<sup>1</sup> Om kontrollmetoden inte kan betraktas som felfri så ska dess standardosäkerhet  $\sigma_{kontroll}^2$  reduceras bort via formeln  $s_{korr} = \sqrt{s^2 - \sigma_{kontroll}^2}$

### Tolerans, standardosäkerhet:

$$s_{Plan} \leq \sigma \cdot \sqrt{F_{0.05,19,\infty}} \approx 20 \cdot (0,96 + (19)^{-0.4}) = 20 \cdot 1.27 = 25.4 \text{ mm}$$

där  $s_{Plan}$  är stickprovets skattade standardosäkerhet. Ev. grova fel från föregående steg ska först tas bort ur materialet .

Vid kontrollen skattades standardosäkerheten till  $s_{Plan} = 21.1$  mm. Eftersom  $21.1 \leq 25.4$  så godkänns det aktuella kontrollområdet i detta avseende.

### Systematik

För ett stickprov på 20 objekt, med den specificerade kvalitén att ingen systematik ska finnas ( $\mu = 0$ ), blir gränsvärdet  $|\bar{x}| \leq s_{Plan} \cdot t / \sqrt{n}$  där  $|\bar{x}|$  är absolutbeloppet av medelavvikelsen. För  $s_{Plan} = 21.1$ ,  $n = 20$  och  $t / \sqrt{20} = 0.39$  (Tabell 5) får vi toleransen

$$\text{Tolerans, systematik: genomsnittlig radiell avvikelse} \\ |\bar{x}| \leq 21.1 \cdot 0.39 = 8.2 \text{ mm.}$$

Vid kontrollen skattades medelavvikelsen till  $\bar{x} = 25.0$  mm. Eftersom  $25.0 > 8.2$  så underkänns hela det aktuella kontrollområdet i detta avseende. Orsaken till den systematiska avvikelsen måste utredas.

## 6 Epilog

### 6.1 Reflektioner från Norgebesöket

Den 10-11 september träffade författarna Håkon Dåsnes, Erling Onstein och Nils Ivar Nes från Kartverket samt Knut Jetlund från Statens vegvesen. Mötet hölls i Hamar där kartverket har ett regionalkontor.

Först några allmänna reflektioner från besöket:

- Inga standarder inom geodataområdet är "Norsk standard". Det är frivilliga branschstandarder utanför den officiella - och dyra - standardiseringsapparaten. Inom FKB finns dock vissa bindande delar i datautväxlingen mellan stat och kommun som baseras på dessa standarder.
- Standardiseringsarbetet och geodatasamarbetet i Norge bygger mer på frivillighet än på styrning, samt på insikten att alla tjänar på att medverka.
- Egentligen skulle branschstandarderna ha reviderats vart tredje år, men så har det av olika skäl inte blivit. Allt arbete ser dock ut att hänga ihop (SOSI, Geodatastandard, Kontroll av geodata etc.).

- Kartverket verkar vara något av motor i det mesta som görs vad gäller geodata, även om det finns en partssammansatt grupp som gör prioriteringar - i samma anda som Geovekst, Norge digital etc.
- Vem som helst kan dock ta initiativ till revideringar, men då förväntas man nästan föreslå sig själv som ansvarig.
- Kartverket har egna grupperingar som jobbar med detta - dels inom GD-stab, dels på Landdivisjonen.
- Håkon Dåsnes, som sitter på Landdivisjonen, är projektledare för översynen av Geodatastandarden och tillhörande dokument. Även de övriga vi träffade finns med i bilden.
- Egentligen skulle arbetet vara klart under detta år, men deadline har förskjutits framåt. Sammantaget ter sig timingen och förutsättningarna ganska goda för ett svensk-norskt samarbete.

Vad gäller specifikt *kvalitetskontroll* kan redovisas:

- Det finns i "Kart og geodata", som hänger intimt samman med "Kontroll av geodata", några gränsvärden som verkar litet gripna ur luften. Dessa passager redovisas i Bilaga 1 till denna rapport, för ett senare ställningstagande; dock hann inte detta diskuteras under besöket.
- Våra synpunkter och idéer rörande detaljutformningen av de olika testerna redovisas i Bilaga 2. De framfördes och diskuterades under besöket och i princip fick vi medhåll. De kommer att ligga till grund för ett fortsatt samarbete.
- Detsamma gäller (som redan tidigare sagts i avsnitt 4.2) terminologin beträffande "toleranser", se Bilaga 3.
- Det är fortfarande oklart i vilken omfattning standarden tillämpas - mot bakgrund av att kontrollerna ter sig väldigt ambitiösa, med stora stickprov och ett omfattande rapporteringskrav. Kartverket gör dock årliga kontroller enligt standarden gentemot ett urval av sina geodataleverantörer.

Avslutningsvis var det mötets samlade mening att tillsammans försöka åstadkomma följande:

- Vi vill gärna - genom samarbete - åstadkomma ett gemensamt koncept för datakvalitet och kvalitetskontroll i Norden /Skandinavien
- Vi vill också försöka sprida detta koncept globalt genom vår medverkan i det internationella standardiseringsarbetet, t.ex. inom ISO.
- Vi bedömer att det som i första hand krävs är språkliga /terminologiska anpassningar.

## 6.2 Slutord

Det ter sig helt rimligt att anpassa kvalitetskontrollen i HMK till den norska standarden "Kontroll av geodata" - åtminstone som metodbeskrivningar. Med det menar vi att kvalitetskontroll av geodata bör göras på detta sätt, utan att det ställs skall-krav på att alltid, i alla sammanhang, genomföra samtliga kontroller.

## 7 Referenser

- Devillers, D. och Jeansoulin, R. (editors), 2006. Fundamentals of spatial data quality. ISTE Ltd., Storbritannien och USA.
- Lantmäteriet, 2012. [HMK-Introduktion, 2012 års arbetsdokument](#). Sverige.
- Lantmäteriverket, 1996. [Handbok till mätningsskuggörelsen - Geodesi, Stommätning](#) (HMK-Ge:S). Sverige
- Statens kartverk, 2001. [Kvalitetssikring av oppmåling, kartlegging og geodata \(Geodatastandard\)](#). 75 s. Norge.
- Statens kartverk, 2007. [Kontroll av geodata](#). 90 s. Norge.
- Statens kartverk, 2009. [Kart og geodata](#). 111 s. Norge.

## Bilaga 1: Eksempel på kvalitetskrav ved kontrollmåting

Føljande konventioner finns i "Kart og geodata". Det finns anledning att ta ställning till i vilken mån var och en av dem ska anammas i HMK. Observera att "krav" och "tolerans" verkar användas synonymt.

- Overstiger restfeilene i transformasjonspunktene 1/3 av kravene til de mest nøyaktige kartobjektene som skal innmåles, må oppdragsgiver kontaktes.
- Ved fri oppstilling skal standardavviket på bestemmelsen av oppstillingspunktet ikke overstige 1/3 av kravene til de mest nøyaktige kartobjektene som skal innmåles.
- Kjentpunktene for en aerotriangulering skal totalt ha et standardavvik maksimalt lik 1/3 av toleransene for geodata som skal produseres fra bildene. Med totalt menes samlet virkning av standardavviket til markmålingen og til definisjonen (bildemåling, signalering, instrumentfeil o.l.) av punktet. Kjentpunktene skal måles inn i marka og bestemmes med et standardavvik maksimalt lik 1/4 av toleransene for geodata.
- Nypunktene skal bestemmes med et gjennomsnittlig estimert standardavvik maksimalt lik 1/2 av toleransene for geodata. De ytre orienteringselementene skal ha et totalt standardavvik maksimalt lik 2/3 av toleransene for geodata.
- De enkelte punktene i kjentflaten skal måles inn med et standardavvik maksimalt lik 1/3 av toleransen for punktene i den ferdige høydemodellen.
- Ved kartkontroll o.l. settes skillet mellom avvik og grove feil for målbare størrelser til en verdi lik 3 ganger toleransen for standardavviket for den aktuelle størrelsen. (Toleransen for standardavviket må være fastsatt i standard eller avtale.) Ved utjevning av nett e.l. konstateres grov feil ved statistisk testing.

## Bilaga 2: Test av grova fel, mätosäkerhet och systematik i 1 D, 2D och 3D

I denna bilaga förs ett resonemang om en nyansering av gränsvärdena i en dimension, två dimensioner (i planet) och tre dimensioner (i rymden) bör införas. Det utgör därigenom även en djupare analys av gränsvärdena i "Kontroll av geodata". Angivna värden har verifierats genom Monte Carlo-simulering.

### Grova fel

$3\sigma$  används som gräns för grova fel. I 1D motsvarar det konfidensgraden 99,73 %. Denna konfidensgrad i två och tre dimensioner ger  $2,43\sigma$  i 2D och  $2,17\sigma$  i 3D. Vill man söka grova fel med samma risknivå så är det alltså följande täckningsfaktorer som ska användas:

1D	$3\sigma$	→	99,73 %
2D	$2,43\sigma$	←	99,73 %
3D	$2,17\sigma$	←	99,73 %

Detta förutsätter dock att koordinatosäkerheterna  $s_x$ ,  $s_y$  respektive  $s_z$  är någorlunda lika stora samt att det inte finns några större korrelationer mellan koordinatvärdena.

Är inte dessa villkor uppfyllda så tenderar täckningsfaktorerna att gå mot 3, dvs. gränsvärdena i 1D-kolumnen ger minst 95 % konfidensgrad även i 2D och 3D. Så vill man hålla risknivån (sannolikheten för att underkänna en korrekt mätning) under 5 % så är det den kolumnen man alltid bör använda.

### Mätosäkerhet

Mätosäkerheten kontrolleras med hjälp av gränsvärden/toleranser beräknade ur F-fördelningen. Även här finns en möjlighet att göra en nyansering av gränsvärdena beroende på dimensionen.

Det sker genom att man ersätter frihetsgraderna  $n-1$  med  $dim*(n-1)$ , där  $dim$  antar värdena 1, 2 och 3 för 1D, 2D respektive 3D.

Detta ger:

n	1D $\sqrt{F_{0.05,n-1,\infty}}$	2D $\sqrt{F_{0.05,2(n-1),\infty}}$	3D $\sqrt{F_{0.05,3(n-1),\infty}}$
5	1,54	1,39	1,32
7	1,45	1,32	1,27
10	1,37	1,27	1,22
15	1,30	1,22	1,18
20	1,26	1,19	1,15
25	1,23	1,17	1,14
35	1,20	1,14	1,11
50	1,16	1,12	1,10
75	1,13	1,09	1,08
100	1,12	1,08	1,07
150	1,09	1,07	1,05
200	1,08	1,06	1,05

Som vi ser ger kolumnen för 1D samma gränsvärden som Tabell 7.2 i "Kontroll av geodata". Minskningen av gränsvärdena för 2D och 3D beror på att sannolikheten för stora koordinatosäkerheter minskar ju fler avvikelser som ingår; att t.ex. x, y och z samtidigt ska anta stora värden är liten.

Förutsättningarna är desamma som vid detektering av grova fel, dvs. ungefär lika stora koordinatosäkerheter och inga nämnvärda korrelationer. Vill man gardera sig mot detta använder man 1D-kolumnen även i 2D och 3D.

## Systematik

Eventuell systematik kontrolleras med hjälp av gränsvärden beräknade ur t-fördelningen. Även här finns en möjlighet att göra en nyansering.

Istället för att fundera på hur en strikt formel skall se ut så har vi gått den enkla vägen och tillämpat Monte Carlo-simulering (100 000 simuleringar).

Detta ger:

<b>n</b>	<b>1D</b> $t_{0,975,n-1}$	<b>2D</b> empiriskt t-värde	<b>3D</b> empiriskt t-värde
5	2,78	2,11	1,86
7	2,45	1,97	1,78
10	2,26	1,88	1,72
15	2,14	1,83	1,68
20	2,09	1,80	1,66
25	2,06	1,78	1,65
35	2,03	1,77	1,64
50	2,01	1,76	1,64
75	1,99	1,75	1,63
100	1,98	1,74	1,62
150	1,98	1,74	1,62
200	1,97	1,74	1,61
$\infty$	1,96	1,73	1,61

I detta fall överensstämmer inte 1D-kolumnen med motsvarande tabell i "Kontroll av geodata" (Tabell 7.3). Det beror på att man där har använt  $t_{0,95}$  i stället för  $t_{0,975}$ . Eftersom t-fördelningen har två "svansar", precis som normalfördelningen, så måste man – som vi ser det – använda  $t_{0,975}$ . Det ger risknivån 2.5 % i varje ände, dvs. totalt  $2 \cdot 2.5 = 5$  %, och konfidensnivån blir följaktligen 95 %. Tabell 7.3 bör därför innebära att konfidensnivån är 90 %, men det intrycket får man inte av texten i "Kontroll av geodata".

Vill man ändå ha kvar sitt synsätt så blir den nyanserade tabellen:

<b>n</b>	<b>1D</b> $t_{0,95,n-1}$	<b>2D</b> empiriskt t-värde	<b>3D</b> empiriskt t-värde
5	2,13	1,77	1,61
7	1,94	1,67	1,56
10	1,83	1,62	1,51
15	1,76	1,59	1,49
20	1,73	1,57	1,48
25	1,71	1,56	1,47
35	1,69	1,55	1,46
50	1,68	1,54	1,46



75	1,67	1,53	1,45
100	1,66	1,52	1,44
150	1,66	1,52	1,44
200	1,65	1,52	1,44
$\infty$	1,64	1,52	1,44

Där överensstämmer 1D-kolumnen med Tabell 7.3 i "Kontroll av geodata".

Förutsättningarna är fortfarande desamma, dvs. om man vill gardera sig mot inhomogen mätosäkerhet och/eller korrelation mellan koordinatvärdena så använder man 1D-kolumnen även i 2D och 3D.

## Bilaga 3: Förslag till harmonisering av terminologin beträffande kvalitetskrav

De följande två bilderna är ett första försök att ”bringa ordning” i terminologin runt kvalitetskrav och toleranser, mot bakgrund av den diskussion som fördes vid Norgebesöket. Den första är på engelska, den andra på svenska – nästa steg borde vara att testa hur detta skulle fungera på norska.

