



Notat: Optimeret N-bestemmelse

31-05-2018

Philipp Trénel, ph.d., seniorkonsulent, specialist
phtr@teknologisk.dk, +45 72 20 33 92

Søren K. Boldsen, ph.d., specialist
sobo@teknologisk.dk, +45 72 20 24 89

Formål

- At bestemme det optimale antal stik-prøver til bestemmelse af en marks N-niveau (i kg N/ha) som funktion af markens størrelse (i ha) og prøvetagningsstrategi {tilfældig sampling, grid sampling, transekt sampling} med en tilsigtet usikkerhed på $u = \{5, 10, 15, 20, 25\}\%$ på baggrund af N-data indsamlet i et 25×25 m grid fra i alt fire danske marker. Opgaven dækker over en genberegning af usikkerhedsbestemmelsen af N-stikprøvetagning i marker jf. proceduren beskrevet i notatet "Optimeret NO_3 -bestemmelsesstrategi" fra den 12-06-2015, se bilag 1, men med to nye marker føjet til.

Overordnede konklusioner

- Udvidelsen af datagrundlaget med to nye marker indikerer, at usikkerheden i bestemmelsen af en marks N-niveau vha. transekt-sampling ikke kun afhænger af antal stikprøver og markens areal, men også varierer fra mark til mark. Dette skyldes formentligt, at marker varierer i forhold til, hvor homogent eller heterogent markerne er i forhold til N-indhold (udtrykt som CV).
- Resultaterne peger dog fortsat på, at transekt-sampling med 16 stik per 5 ha resulterer i en usikkerhed, der enten er signifikant mindre end 10% (gældende for de mere homogene marker) eller ikke kan afvises at være lig 10% (gældende for de mere heterogene marker 49-0 og 78-0). Med 17 stik pr. 5 ha kan det for alle marker ikke afvises, at usikkerheden er $\leq 10\%$.
- Der er fortsat behov for at udvide mark-sættet for at øge generaliserbarheden af konklusionerne.

Metoder

- For beskrivelsen af metoderne, se bilag 1.
- De to nye marker, 4-0 og 78-0, fremgår af Fig. 1.
- Tabel 1 viser en sammenligning af markerne.
- Det samlede sæt marker fremgår af Fig. 2.
- Alle analyser er udført i R (R 2018).

Tabel 1: Marker, anvendt i indeværende analyse.

Mark	Areal (ha)	N-niveau (kgN/ha)	N-variation (CV i %)
4-0	4.9 (lille)	59.6 (høj)	39.8
49-0	9.9	38.5	49.6 (høj)
41-0	9.5	33.6	37.5
78-0	8.5	16.6 (lav)	46.4 (høj)

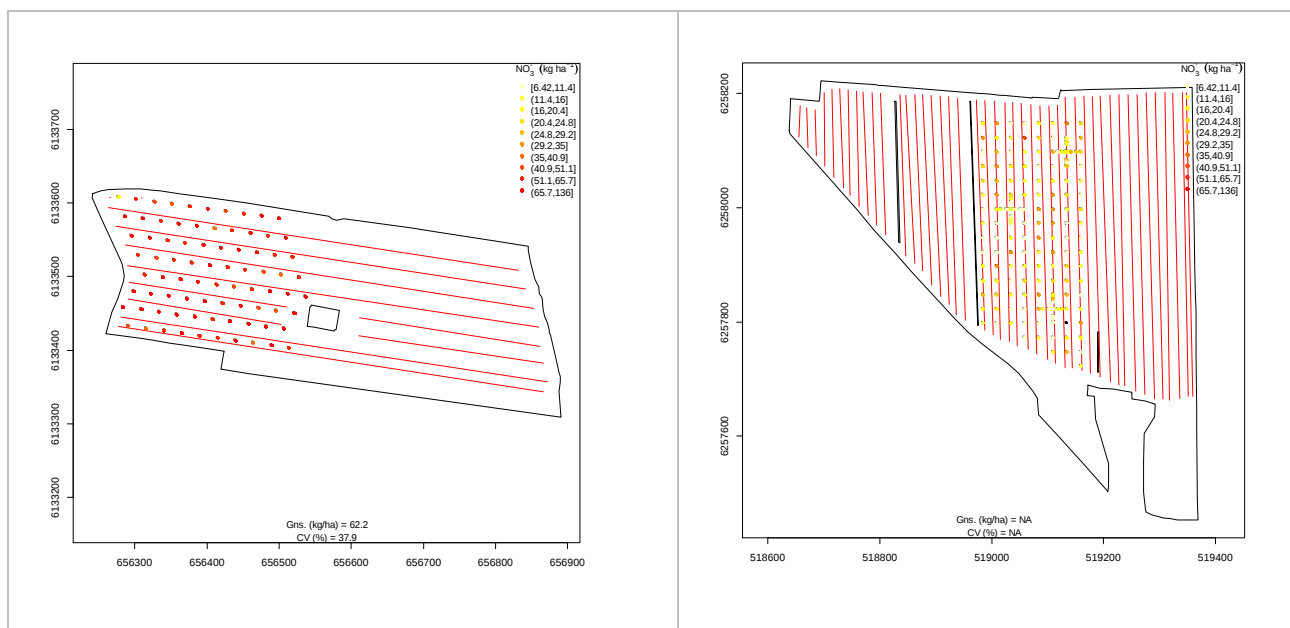


Fig. 1: N-stikprøver i 25 × 25 m grid for Mark 4-0 (til venstre) og Mark 78-0 (til højre).

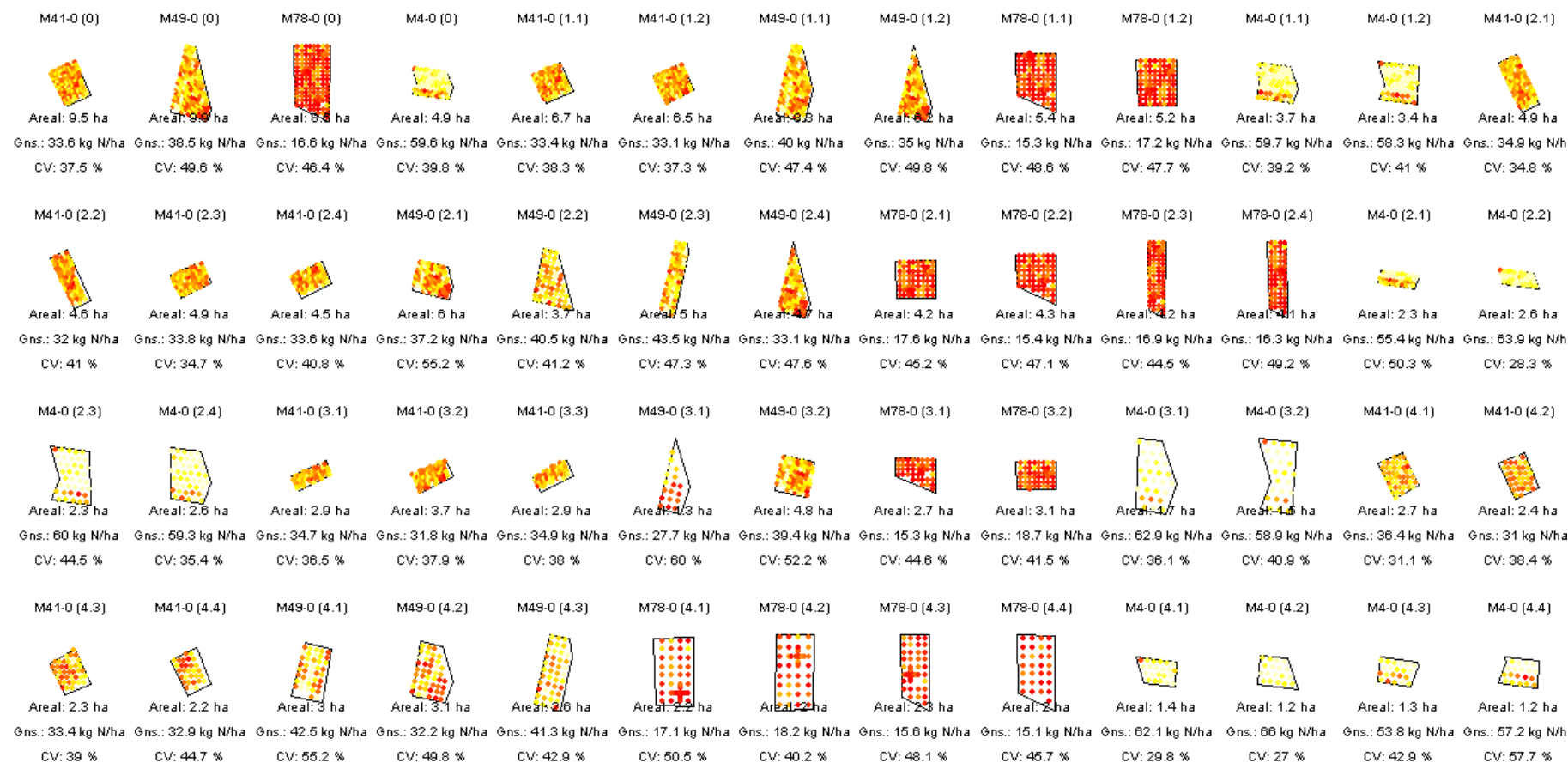


Fig. 2: Sættet af markerne og tilhørende stikprøvesamples brugt i indeværende analyse. M41-0 (0), M49-0 (0), M78-0 (0) og M4-0 (0) er de oprindelige marker; de resterende marker er simulerede marker ved hjælp af opdeling af de to oprindelige marker.



Resultater

- Resultater fremgår af Fig. 3 til 7 og tabel 2 og 3.
- Transekt-sampling udviser tendentielt samme faldende usikkerhed med stigende antal stik i alle marker og delmarker, dog undtagen i mark 78-0 og i dets delmarker, se Fig. 3.
- Den statistiske analyse af bestemmelsesusikkerheden $\log(\widehat{SD(N03^-)})$ udtrykt i % af markens gns. $N03^-$ som funktion af markens areal, antal stikprøver, samplingsstrategien og de fire udgangsmarker (Mark) under hensyntagen af en tilfældig mark-til-mark-variation viste, at Mark er en signifikant term i modellen, se tabel 2; $R^2 = 0.85$.
- Inspektion af modellens resultater indikerer, at det er mark 49-0 og 78-0, som udviser en større usikkerhed, se tabel 3. Begge marker har en høj N-variation eller er mere patchy end de andre to marker, se Tabel 1 og Fig. 2.
- Tabel 3 viser, at den estimerede usikkerhed ved 16 stikprøver pr. 5 ha er signifikant mindre end 10% (M4-0, M41-0) eller ikke kan afvises at være på 10% (M78-0, M49-0). Med 17 stikprøver pr. 5 ha udviser alle marker en usikkerhed, der ikke kan afvises at være $\leq 10\%$.

Tabel 2: ANOVA tabel for en statistisk analyse af bestemmelsesusikkerheden $\log(\widehat{SD(N03^-)})$ udtrykt i % af markens gns. $N03^-$ som funktion af markens areal, antal stikprøver, samplingsstrategien og de fire udgangsmarker (Mark) under hensyntagen af en tilfældig mark-til-mark-variationen.

	numDF	denDF	F-value	p-value
(Intercept)	1	5021	11079.662	<.0001
sqrtN	1	5021	23869.928	<.0001
sqrtAreal_ha	1	44	281.506	<.0001
strategi	2	5021	158.035	<.0001
Mark	3	44	17.218	<.0001
sqrtN:sqrtAreal_ha	1	5021	1280.725	<.0001
sqrtN:strategi	2	5021	20.091	<.0001
sqrtN:Mark	3	5021	15.914	<.0001
sqrtAreal_ha:strategi	2	5021	7.873	0.0004
sqrtAreal_ha:Mark	3	44	3.249	0.0306
strategi:Mark	6	5021	10.416	<.0001
sqrtN:sqrtAreal_ha:strategi	2	5021	2.063	0.1272
sqrtN:sqrtAreal_ha:Mark	3	5021	28.178	<.0001
sqrtN:strategi:Mark	6	5021	2.378	0.0269
sqrtAreal_ha:strategi:Mark	6	5021	3.186	0.0040
sqrtN:sqrtAreal_ha:strategi:Mark	6	5021	1.766	0.1019

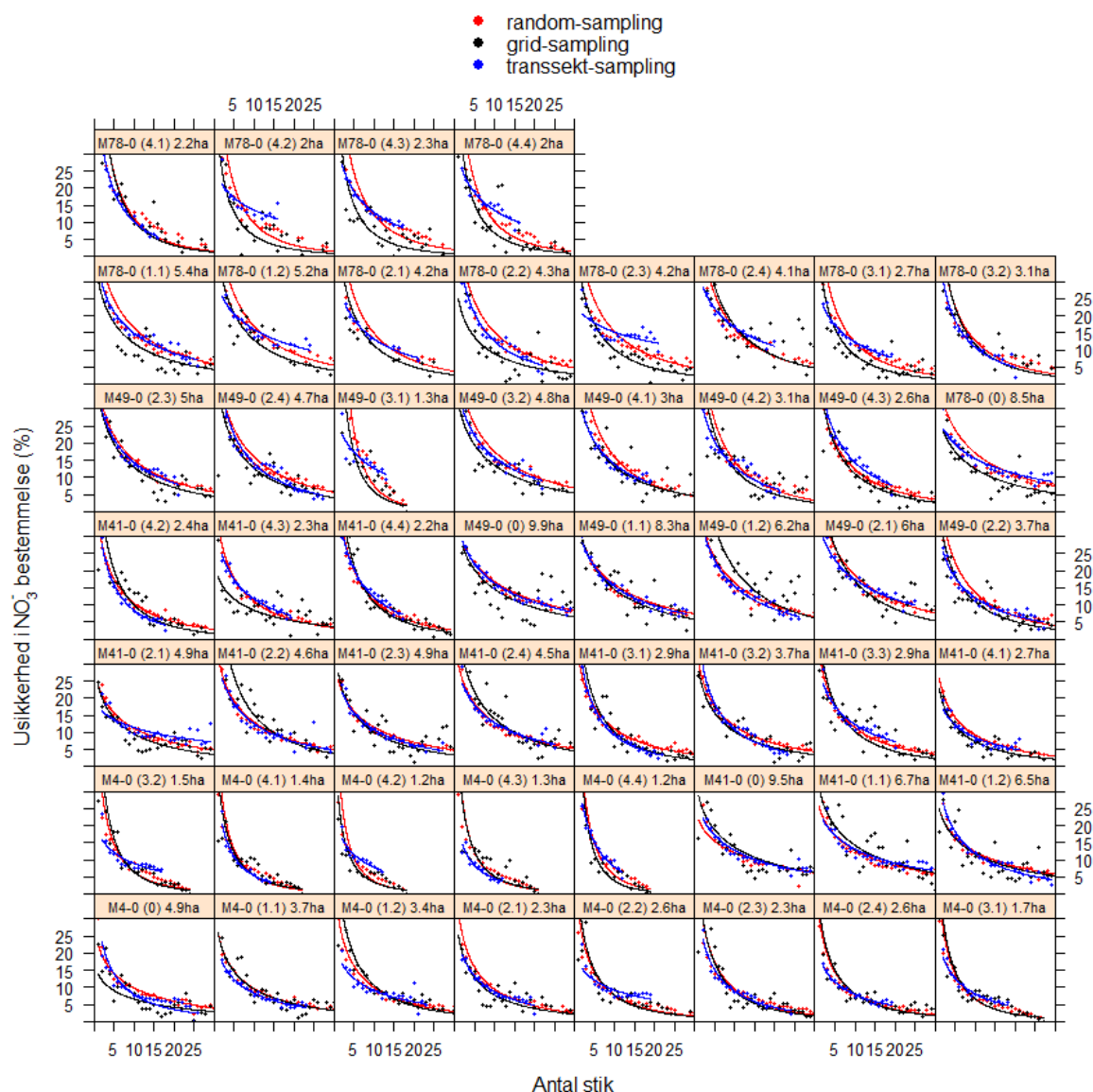


Fig. 3: N-bestemmelsesusikkerhed ($\text{SD}(\widehat{\text{NO}_3^-})$) udtrykt i % af markens gns. NO_3^- som funktion af antal stikprøver for hhv. grid sampling strategien, tilfældig sampling og transekt sampling (for alle marker). Linjer viser for hvert panel den modellerede funktion $\log(\text{SD}(\widehat{\text{NO}_3^-})) = \beta_0 + \beta_1 \sqrt{\text{antal stik}} + e$, hvor $e \sim N(0, \sigma^2)$.

Tabel 3: Estimerede usikkerheder i % ved 16 hhv. 17 stikprøver pr. 5 ha for de forskellige marker M4-0, M41-0, M49-0 og M78-0. P-værdien angivet i de sidste to kolonner er den tosidede hhv. ensidede testsandsynlighed for at usikkerheden u er = 10% hhv. $\leq 10\%$.

strategi = random-sampling:

Antal stik	Areal(ha)	Mark	Usikkerhed (%)	nedre	95%	øvre	95%	.	Pr(u=10)	Pr(u≤10)
16	5	M4-0	8.8	7.3	10.6	ab	0.165	0.918		
16	5	M41-0	8.5	7.9	9.3	a	0.000	1.000		
16	5	M49-0	10.7	9.9	11.6	b	0.090	0.045		
16	5	M78-0	10.4	9.4	11.4	b	0.433	0.216		

strategi = grid-sampling:

Antal stik	Areal(ha)	Mark	Usikkerhed (%)	nedre	95%	øvre	95%	.	Pr(u=10)	Pr(u≤10)
16	5	M4-0	7.3	6.1	8.8	a	0.002	0.999		
16	5	M41-0	7.9	7.3	8.6	a	0.000	1.000		
16	5	M49-0	9.0	8.3	9.8	a	0.013	0.994		
16	5	M78-0	7.6	6.9	8.4	a	0.000	1.000		

strategi = transsekt-sampling:

Antal stik	Areal(ha)	Mark	Usikkerhed (%)	nedre	95%	øvre	95%	.	Pr(u=10)	Pr(u≤10)
16	5	M4-0	7.3	6.0	9.0	a	0.003	0.998		
16	5	M41-0	8.3	7.6	9.1	a	0.000	1.000		
16	5	M49-0	10.3	9.4	11.3	b	0.566	0.283		
16	5	M78-0	10.9	9.8	12.2	b	0.124	0.062		

strategi = random-sampling:

Antal stik	Areal(ha)	Mark	Usikkerhed (%)	nedre	95%	øvre	95%	.	Pr(u=10)	Pr(u≤10)
17	5	M4-0	8.4	7.0	10.1	ab	0.064	0.968		
17	5	M41-0	8.1	7.5	8.8	a	0.000	1.000		
17	5	M49-0	10.2	9.4	11.0	b	0.682	0.341		
17	5	M78-0	9.8	8.9	10.8	b	0.723	0.638		

strategi = grid-sampling:

Antal stik	Areal(ha)	Mark	Usikkerhed (%)	nedre	95%	øvre	95%	.	Pr(u=10)	Pr(u≤10)
17	5	M4-0	7.0	5.8	8.4	a	0.000	1.000		
17	5	M41-0	7.5	6.9	8.1	a	0.000	1.000		
17	5	M49-0	8.5	7.8	9.2	a	0.000	1.000		
17	5	M78-0	7.2	6.5	7.9	a	0.000	1.000		

strategi = transsekt-sampling:

Antal stik	Areal(ha)	Mark	Usikkerhed (%)	nedre	95%	øvre	95%	.	Pr(u=10)	Pr(u≤10)
17	5	M4-0	7.1	5.8	8.7	a	0.001	0.999		
17	5	M41-0	7.9	7.2	8.6	a	0.000	1.000		
17	5	M49-0	9.8	8.9	10.7	b	0.595	0.702		
17	5	M78-0	10.5	9.3	11.7	b	0.432	0.216		

Confidence level used: 0.95

P value adjustment: tukey method for comparing a family of 4 estimates

significance level used: alpha = 0.05

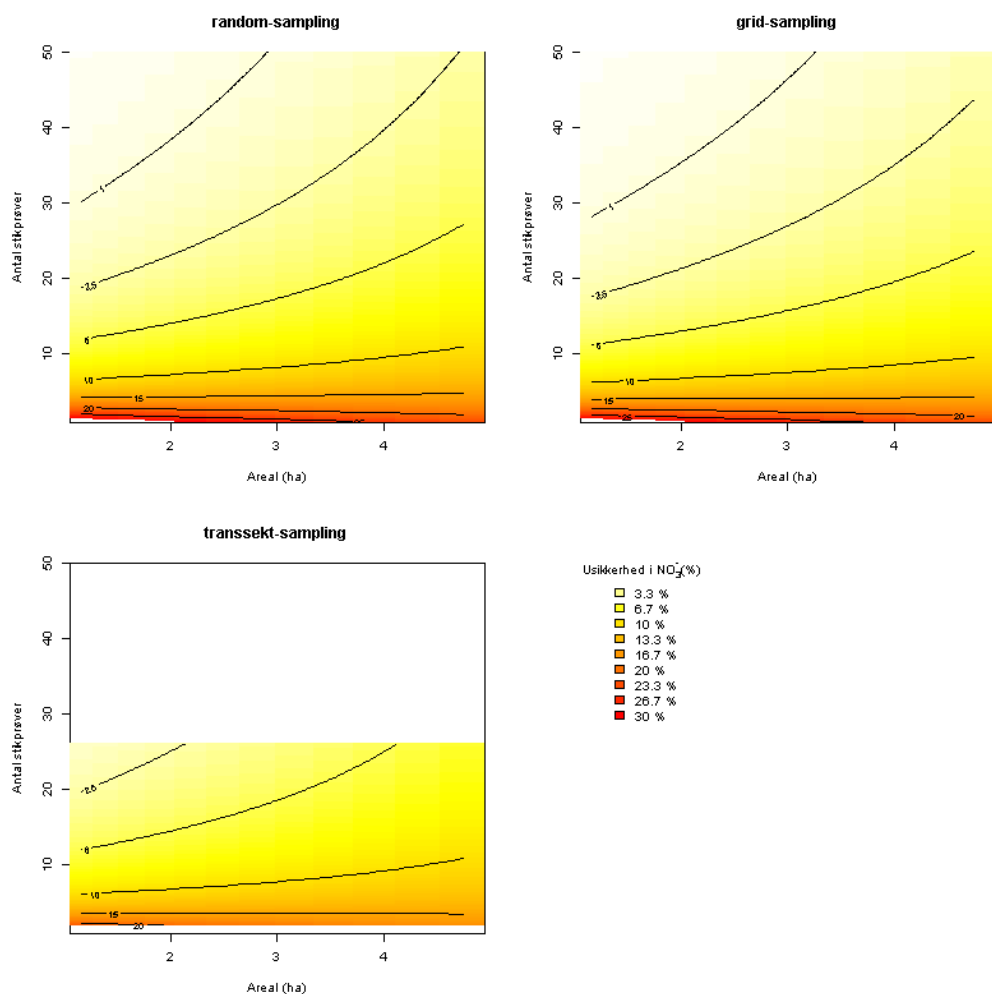


Fig. 4: For mark **4-0** antal stikprøver som funktion af markens areal (ha) ved stigende NO_3^- bestemmelsesusikkerheder ($\text{SD}(\overline{\text{NO}_3^-})$) udtrykt i % af markens gns. NO_3^- for hhv. grid sampling strategien, tilfældig sampling og transekt sampling.

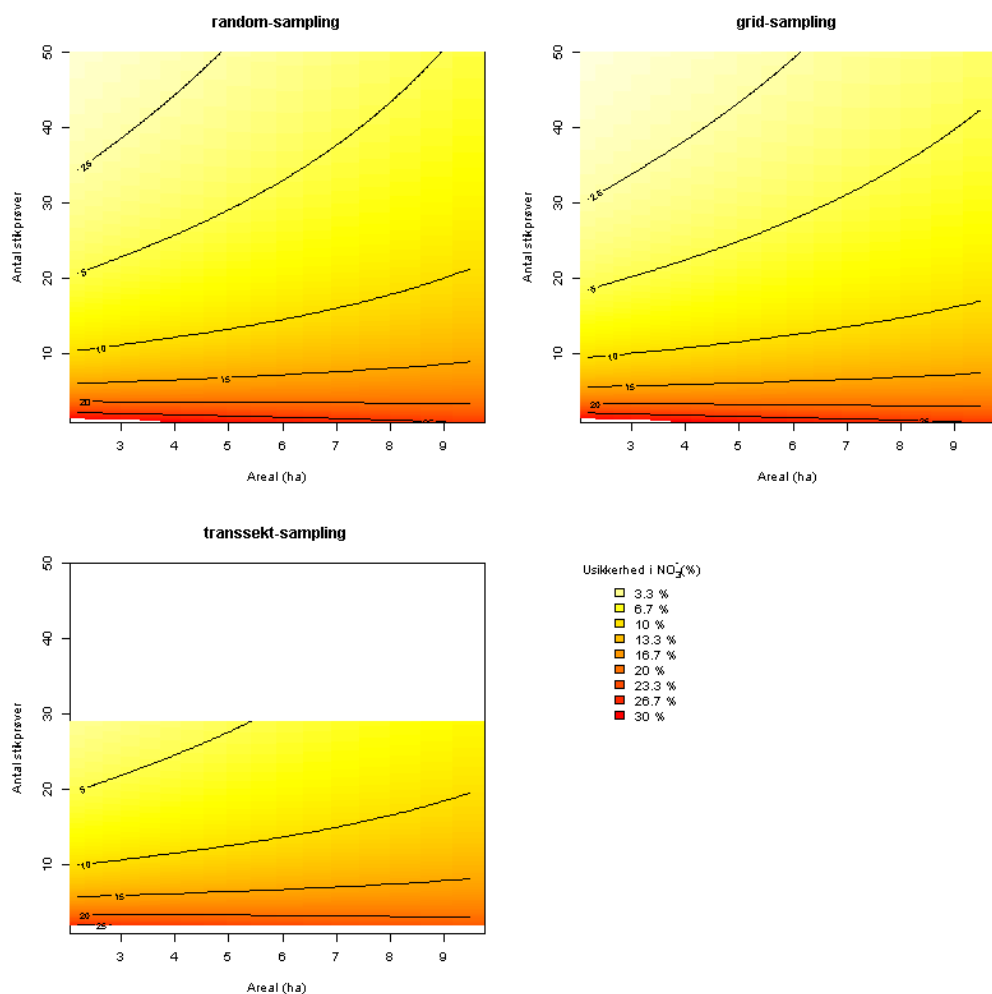


Fig. 5: For mark **41-0** antal stikprøver som funktion af markens areal (ha) ved stigende NO_3^- bestemmelsesusikkerheder ($\text{SD}(\widehat{\text{NO}_3^-})$) udtrykt i % af markens gns. NO_3^- for hhv. grid sampling strategien, tilfældig sampling og transekt sampling.

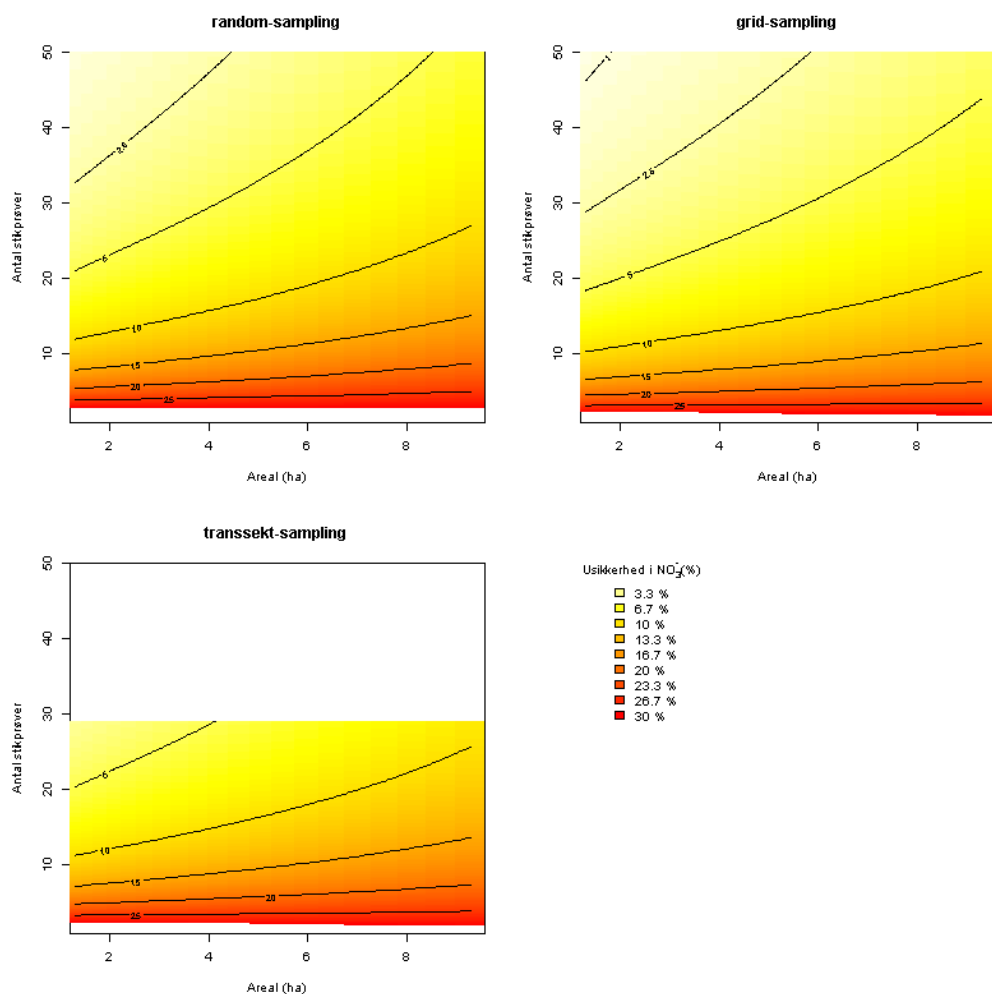


Fig. 6: For mark **49-0** antal stikprøver som funktion af markens areal (ha) ved stigende NO_3^- bestemmelsesusikkerheder ($\text{SD}(\widehat{\text{NO}_3^-})$) udtrykt i % af markens gns. NO_3^- for hhv. grid sampling strategien, tilfældig sampling og transekt sampling.

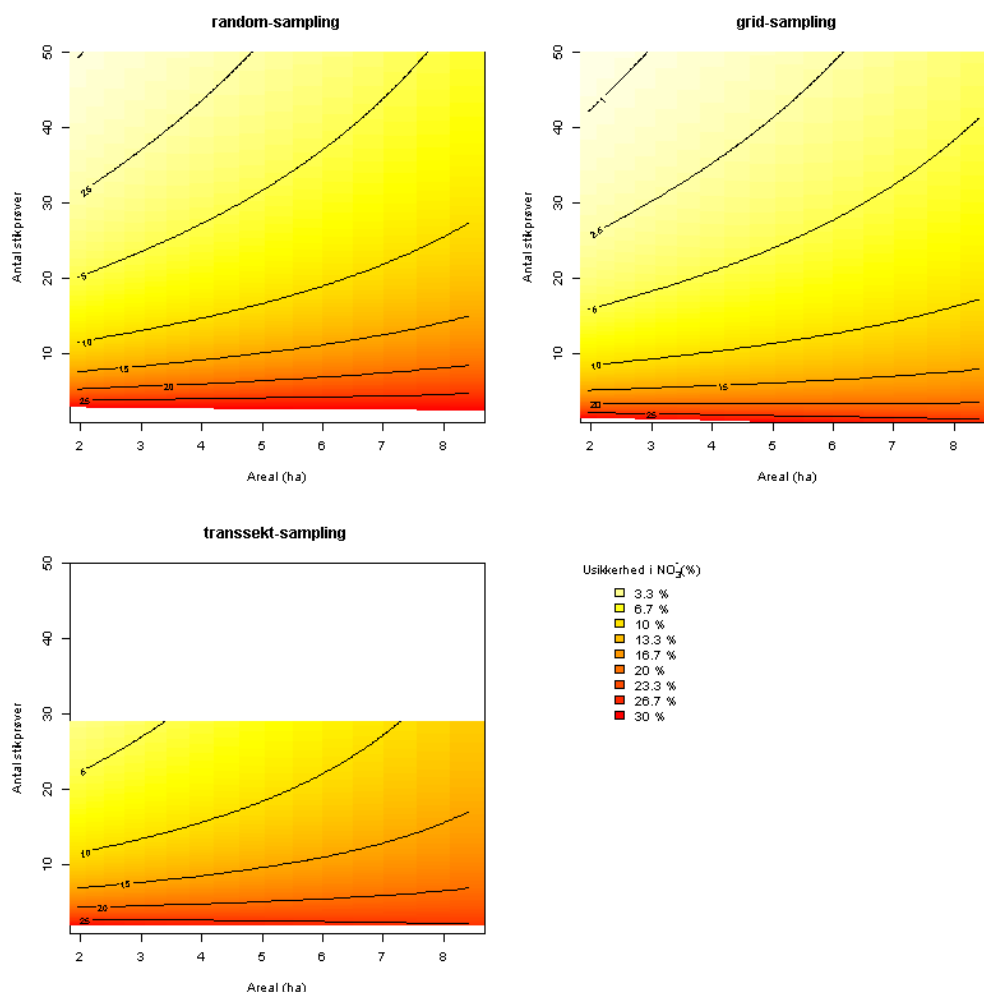


Fig. 7: For mark **78-0** antal stikprøver som funktion af markens areal (ha) ved stigende NO_3^- bestemmelsesusikkerheder ($\text{SD}(\widehat{\text{NO}_3^-})$) udtrykt i % af markens gns. NO_3^- for hhv. grid sampling strategien, tilfældig sampling og transekt sampling.

Referencer

R Development Core Team (2018). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL <http://www.R-project.org>.