Norrbyn

1. Objectives

The present work includes the following analyses:

- An investigation into the consequences for main results of reducing the sample size presently used in monitoring this area.
- An investigation into the consequences for main results of adopting a different (optimal) sample allocation scheme across strata than the one presently used.
- An investigation into the consequences for main results of changing both the sample size and the sample allocation scheme used in monitoring this area.
- An investigation into the consequences for the trends of main indicators of sampling every two years and every three years

2. Methodological notes

2.1. Fixed stations and Stratification

The sampling design includes a set of fixed stations that are sampled annually without replacement. For simplicity, the estimates of variance developed in the present study assume the data was collected using simple random sampling with replacement.

Additionally, the present sampling design involves a stratification of the area into four depth strata (0-3m; 3-6m; 6-10m; 10-20m). To date results at area level have been computed using a simple average of the values of each indicator across all samples collected in the year. This procedure ignores the underlying stratification used in collecting the sample and leads to biased estimates when the samples are combined to produce area-level results[1]. Correcting this bias require the weight of the different strata (e.g., their area or volume) to be known and included in the estimation procedure. At the time of the analysis such weights were not available. To alleaviate the impact of previous effect on analyses and keep results comparable accross sample sizes and allocation schemes an assumption was made that the sampling effort deployed in each stratum was proportional to the size of the stratum. This assumption is likely not valid for the area analysed so results of the present study should be regarded with particular caution until the issue is definitively corrected.

2.2. Resampling strategy, power and interpretation of results

This study carries out a series of re-sampling analyses that provide insights into the consequences of sample size reductions and re-distribution of effort across strata for the precision obtained in a set of statistical indicators, namely their mean value. Albeit providing

[1] Note that the strata-level estimates are not affected by this bias; and that, as long as the number of stations in each strata is maintained across years (a situation that in fact happened throughout the time series), the trends in mean values are still worth analyzing. The results of the optimizations, however, will conditional on the maintanence of the strata weights used in the original (and present) sample allocation.

significant insight into those effects and approximating the reasoning behind a power analysis, these analyses are not *de facto* power analyses. As a consequence, the results obtained with regards to sample sizes and re-allocation of effort should not be interpreted as providing the sampling levels or strategies required for sufficient detection of particular changes in the system or the testing of hypotheses; rather they indicate sample size reductions and re-distributions of effort that, based on available data, yield approximately the same results as obtained by the current programme irrespective of the power they provide.

2.3. Resampling and re-allocation

Re-sampling of annual estimates: 5000 Bootstrap replicates of area-level and strata-level annual estimates were generated. These were done using simple random sampling with replacement and the original sample sizes per strata as stratum weight. Different sample sizes and sampling effort allocations were tested, including single indicator Neyman allocation and compromise multi-indicator Neyman allocation. In compromise allocations, the minimum number of stations accepted for each stratum was 5 with re-allocations being made across the remaining strata when expected sample sizes from Neyman allocation were below this threshold. In both single indicator- and compromise-allocation, when a strata is not presently used to calculate the average for a certain indicator its optimal sample size was considered 0.

Re-sampling of trend estimates: 1000 bootstrap replicates of regressions of mean indicator ~ year were determined. When simulating sampling every two years a random start for the series was defined, with the first sampled year being either the 1st or the 2nd year available in the time series. When simulating sampling every three years random start for the series was also defined, with the first sampled year being selected among the 1st, or 2nd, or 3rd year available in the time series. To keep the length of the time series constant and secure comparability of results a similar limitation was put on the end year used in analysis. E.g., in the case of simulations of sampling every two years involving the 1st year available, every second year was included until year (t-1), i.e., the second last year in the series; in the case simulations that involved starting in the 2nd year available, every second year was included until year (t), i.e., the last in the series. A similar reasoning was used in simulations of sampling every three years.

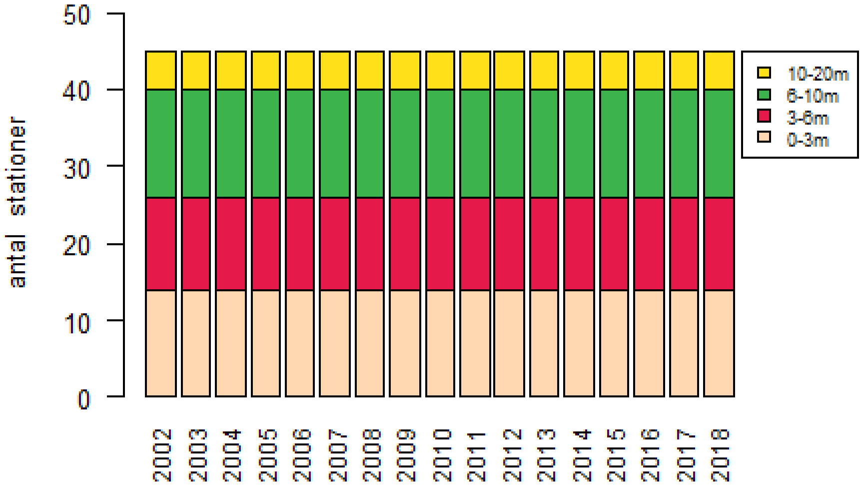
2.4. Data available

The data consisted of the numerical values for the indicators CodN, CyprinidsB, CyprinidsN, FlounderN, HerringN, PerchB, PerchN, PikeN, PikeperchN, PiscivoresN, and WhitefishN by station, **from 2002 to 2018**, as defined in the following table

Indicator	Swedish	Definition	Strata used in calculating average
CodN	CPUE Torsk	Cod, Number per gear	"0-3m", "3-6m", "6-10m", "10-20m"
CyprinidsB	WPUE Karpfisk	Cyprinid fish, Biomass per gear	"0-3m", "3-6m", "6-10m"
CyprinidsN	CPUE Karpfisk	Cyprinid fish, Number per gear	"0-3m", "3-6m", "6-10m"
FlounderN	CPUE Skrubbskädda	Flounder, Number per gear	"0-3m", "3-6m", "6-10m", "10-20m"
HerringN	CPUE Strömming	Herring, Number per gear	"0-3m", "3-6m", "6-10m", "10-20m"
PerchB	WPUE Abborre	Perch, Biomass per gear	"0-3m", "3-6m", "6-10m"
PerchN	CPUE Abborre	Perch, Number per gear	"0-3m", "3-6m", "6-10m"

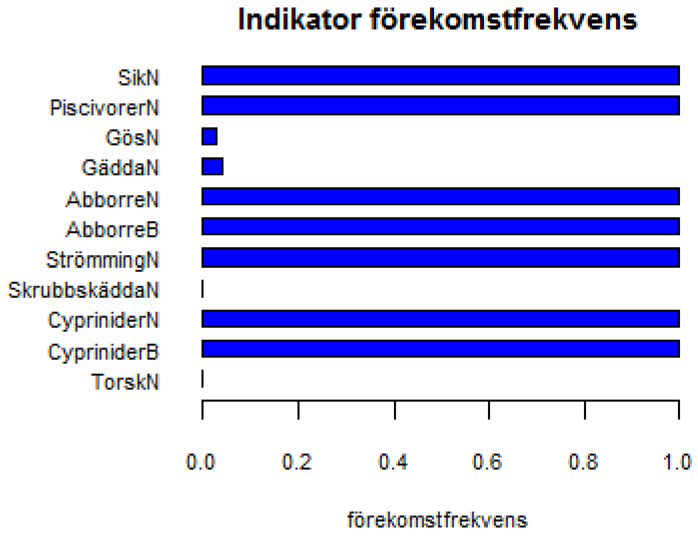
PikeN	CPUE Gädda	Pike, Number per gear	"0-3m", "3-6m", "6-10m"
PikeperchN	CPUE Gös	Pikeperch, Number per gear	"0-3m", "3-6m", "6-10m", "10-20m"
PiscivoresN	CPUE Rovfisk	Piscivorous fish, Number per gear	"0-3m", "3-6m", "6-10m"
WhitefishN	CPUE Sik	Whitefish, Number per gear	"0-3m", "3-6m", "6-10m", "10-20m"

The number of stations sampled **per year (n=45),** their distribution across strata, and the methodology used during sampling were constant throughout the time series. All stations were sampled every year apart from some departures due to disturbances in the fishing area (**54 out of 765 stations sampled**).



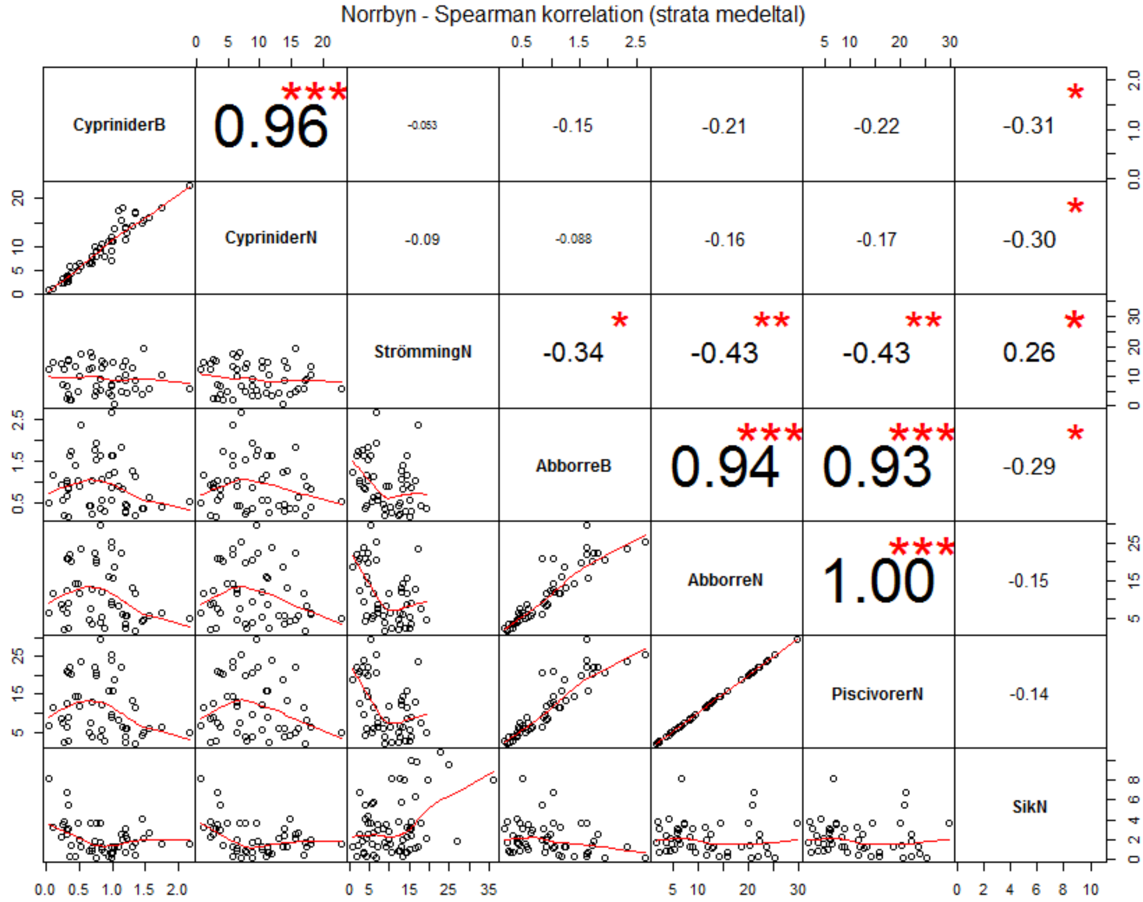
2.5. Choice of indicators

Results were obtained for all indicators in sample size and sample allocation scenarios tested. However, only the ones derived for species that register higher frequency of occurrence in the area were considered when defining the optimum scenarios for re-allocation of stations across depth strata. This is because it is difficult to obtain precise estimates for rare and less common species that register a large number of zero-observations unless a dedicated programme is established that specifically targets the habitats (e.g., depths) where they exist. **In the case of Norrbyn, these were the Cyprinids, Herring, Perch, Piscivores and Whitefish.**



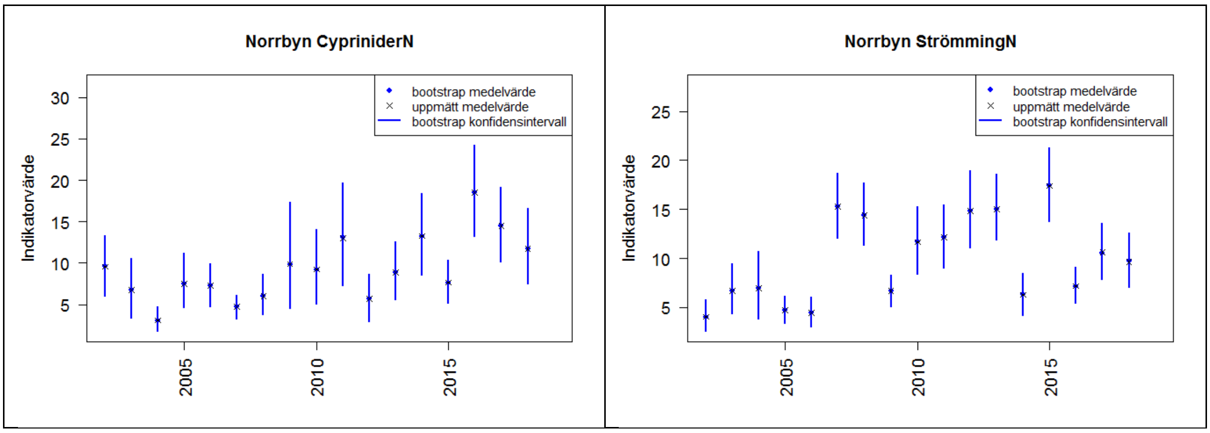
Furthermore, during initial analyses, some indicators were identified as highly positively correlated with each other (e.g., CyprinidsB and CyprinidsN). The presence in the analysis of indicators with very high and significant positive correlations is not particularly informative on the status of the system (the indicators are likely to reflect the same pattern) and has the negative effect of giving them excessive weight in the results of the allocation algorithm (thus making the results less optimal for other indicators, particularly those with contrasting distributions). It was therefore considered useful to further restrict the indicators used in studies of re-distribution of samples across depth strata to the subset not displaying such correlations.

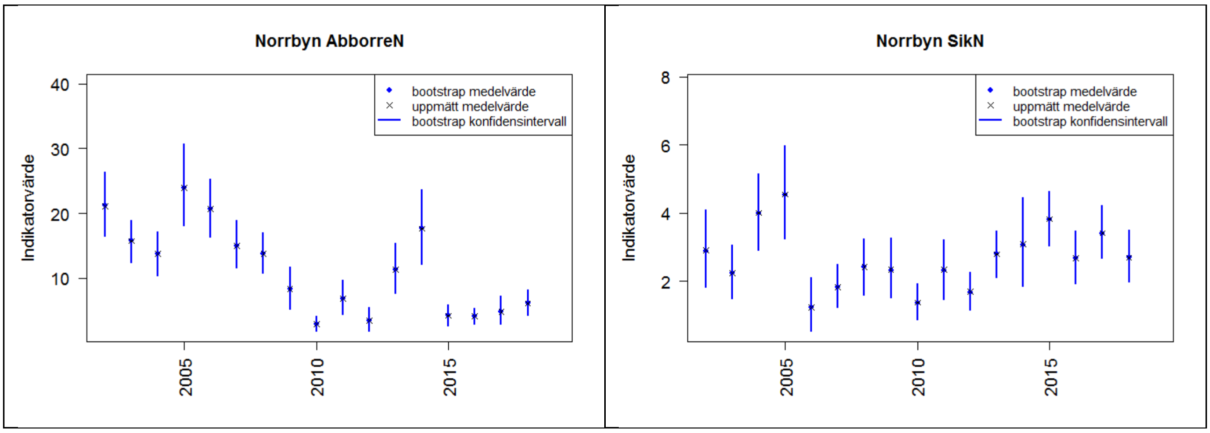
In the case of Norrbyn, high positive significant correlations are observed between CyprinidsN and CyprinidsB, and between PerchN, PerchB and PiscivoresN. After these redundancies were eliminated, **CyprinidsN, HerringN, PerchN and WhitefishN** remained as the main indicators to be used in the analyses.



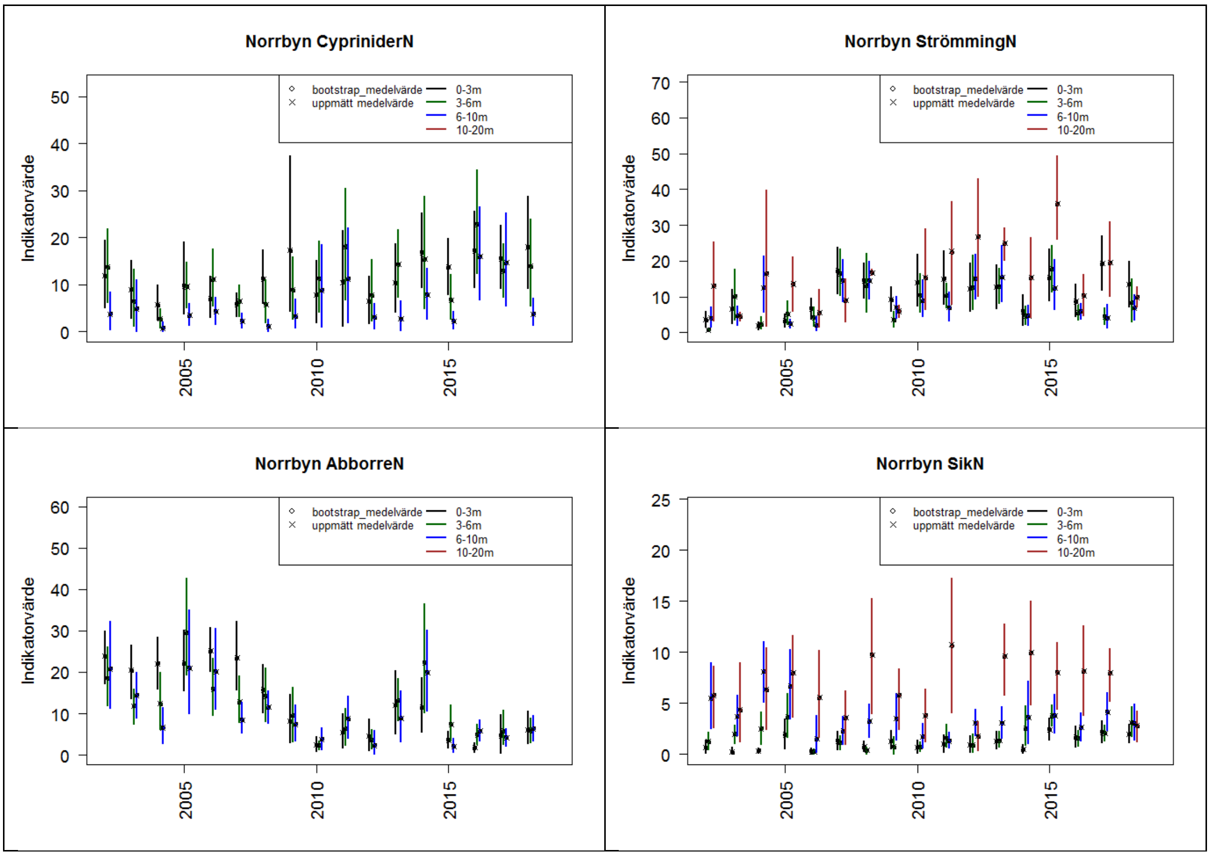
3. Results

3.1. Variation in indicator over the years



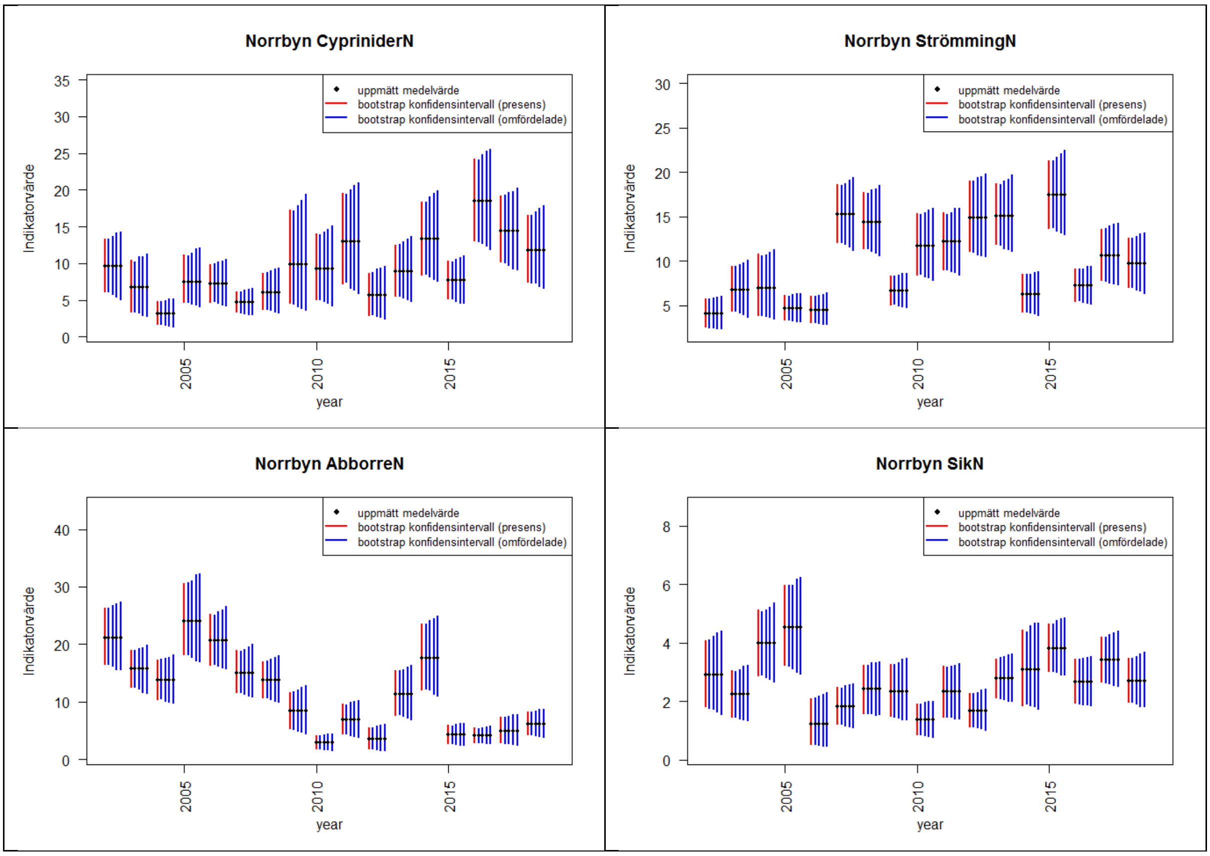


3.2. Variation in indicator over the years by depth strata



3.3. Variability in results with sample size (original allocation)

The figure displays the impacts of sample size reductions in present area-level estimates under the present sampling **effort of 45 stations** (red and first blue line in each series) and under successively smaller sampling **effort of 39, 34 and 31 stations** (remainder blue lines, from left to right within each year) [2]. The simulations were quite stable as shown by the low variability in the results of the two first confidence intervals of each year (compare red and first blue line; first two rows of table). The decrease in precision that a reduction in sample size could have caused can be observed in the relative increase of the confidence intervals from left to right within each year.



The following table displays detailed results on the 5%, 50% (median) and 95% quantiles of the distribution of relative standard errors (RSE) of the simulated replicates. Green coloured cells are estimates that stayed within +5% of the presently obtained value. Red coloured cells contain estimates that are beyond that limit. The comparison of the first two rows provides insight into the variability brought about by the simulations themselves. The increase in RSE observed with decreasing sample size provides insight into the decrease in the precision of area-level estimates to be expected from a reduction in sample size.

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Present	15.7	20.4	27.9	11.1	13.9	21.8	10.8	16.5	24.4	11.4	17.3	23.5
45	15.6	20.5	28.1	11.0	14.2	21.8	10.7	16.2	24.8	11.7	17.6	23.0
39	17.2	22.1	30.5	12.0	15.3	22.2	11.6	17.5	26.4	12.4	18.7	24.7
34	18.4	24.0	32.9	12.8	16.4	23.6	12.8	19.1	28.4	13.3	20.1	26.4
31	19.6	25.4	35.0	13.8	16.8	25.4	13.4	19.9	29.7	13.9	20.4	27.7

[2] The variability in samples sizes tested in the different scenarios considered for this area results from the need to maintain at least 5 samples in all strata while avoiding substantial departures from the strata weights determined for each scenario.

25	22.0	29.6	40.5	15.5	19.5	29.3	15.5	22.9	34.5	15.6	21.9	31.1

Based on these results it can be concluded that **if 39 or 34 stations** had been sampled, the relative standard error (RSE) obtained for the main indicators would most likely have stayed within a +5% interval of present value. This reduction in sampling would correspond to the following re-allocation of stations across strata (changes to weight of strata highlighted in parenthesis):

Depth strata	Present (n=45)		Reduction to n = 39		Reduction to n = 34	
0-3 m	14	(.31)	12	(0,31)	10	(0,29)
3-6 m	12	(.27)	10	(0,26)	9	(0,26)
6-10 m	14	(.31)	12	(0,31)	10	(0,29)
10-20 m	5	(.11)	5	(0,13)	5	(0,15)

3.4. Variability in results with sample size (Neyman allocation)

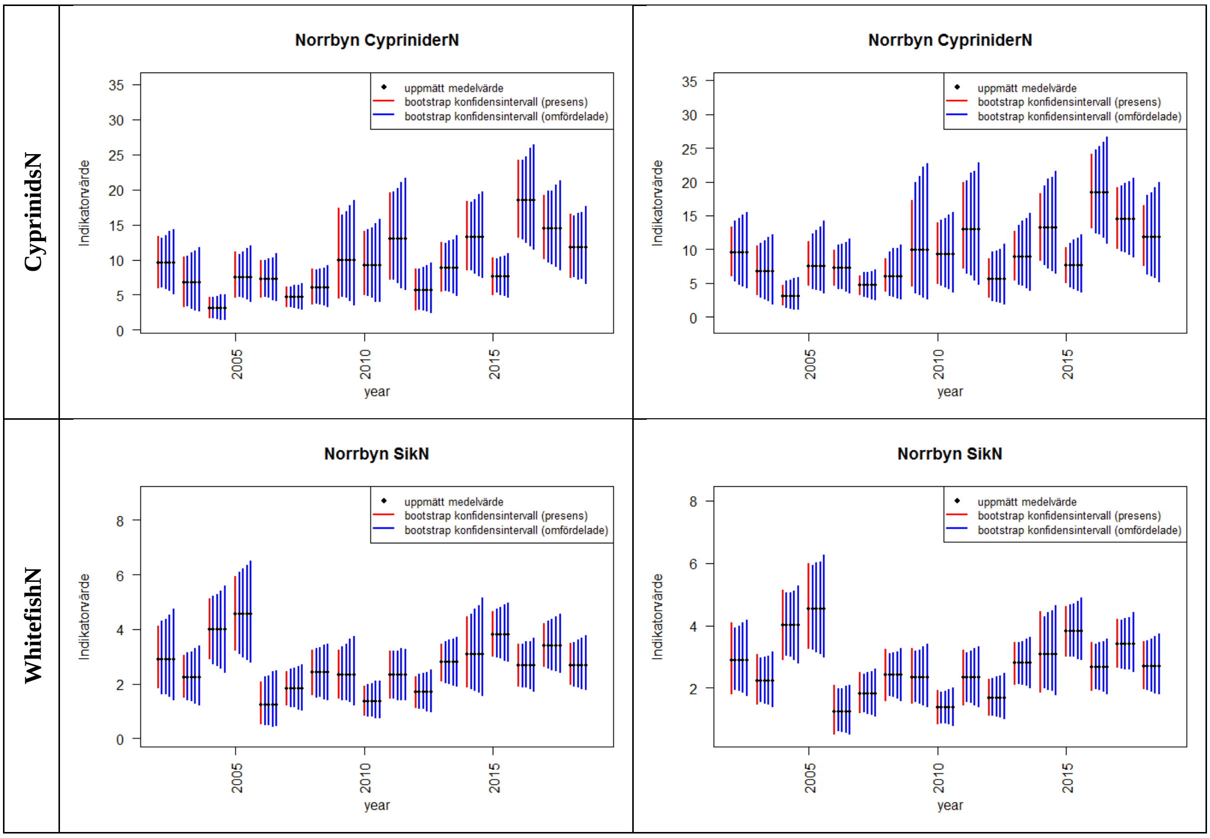
3.4.1. Single Indicator Neyman allocation

The redistribution of sampling effort across strata as indicated by Neyman allocation focused on improving area-level estimates of each of the main indicators is displayed in the following table

Depth strata	Present	CyprinidsN	HerringN	PerchN	WhitefishN
0-3 m	14	17	16	14	8
3-6 m	12	13	10	12	8
6-10 m	14	10	13	14	20
10-20 m	5	5	6	5	9

The following graphs display the evolution of the simulated confidence intervals of two contrasting indicators (rows) under two contrasting allocation scenarios (columns). Each graph displays the confidence interval of the original series (red line) and confidence intervals obtained with successively smaller sample sizes (blue lines). Full results for all scenarios and indicators are displayed in the table that follows. In this table values are expressed in terms of relative standard error (RSE) as calculated from bootstrap. To facilitate interpretation a colour code is used in the cells – Yellow when values are lower than those presently obtained (first row); Green when RSE are within +5% of present values; and red when RSE values are beyond that 5% of the present value.

	Reallocation Scenario	
	CyprinidsN	WhitefishN



The results show that the adoption of a scheme focused on the optimal allocation of one indicator results in more precise estimates for that indicator but frequently generates a negative side-effect on other indicators, which precision significantly degrades relative to its original values. These effects largely motivated the need to consider compromise multi-indicator allocations such as the ones proposed in section 3.4.2.

Reallocation Scenario: focus on CyprinidsN

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Present	15.6	20.4	28.0	11.1	14.1	21.6	11.0	16.3	24.2	11.5	17.3	23.3
45	15.8	19.4	27.9	11.2	14.0	23.3	11.3	16.4	23.9	12.7	18.8	26.1
40	17.0	20.8	30.2	12.2	14.5	23.2	12.2	18.1	25.3	13.4	19.4	27.4
35	18.6	22.5	32.9	13.0	16.0	25.0	13.4	19.3	28.2	14.6	20.6	30.0
30	20.5	25.2	35.8	14.5	17.2	26.7	14.5	21.4	29.8	15.5	21.8	32.7
25	22.8	27.9	40.5	15.9	19.1	28.7	16.1	23.6	34.1	17.3	23.6	35.5

Reallocation Scenario: focus on HerringN

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Present	15.6	20.4	28.1	11.2	14.1	21.8	10.8	16.5	24.3	11.3	17.2	23.6
45	16.1	20.8	28.2	11.2	14.1	21.0	10.9	16.4	24.7	11.6	17.6	23.2
40	17.2	22.3	30.4	12.0	14.9	22.0	11.6	18.0	26.0	12.3	18.2	24.7
35	18.3	23.1	32.4	12.9	15.9	23.6	12.8	19.1	27.3	13.1	19.7	26.8
30	’---	’---	’---	’---	’---	’---	’---	’---	’---	’---	’---	’---

25	'---	'---	'---	'---	'---	'---	'---	'---	'---	'---	'---	'---

Reallocation Scenario: focus on PerchN

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Present	15.5	20.5	28.0	11.1	14.1	21.5	11.0	16.3	24.2	11.4	17.2	23.3
45	15.8	20.5	28.1	11.1	14.4	21.9	10.7	16.1	24.5	11.7	17.6	22.9
40	17.0	21.9	30.7	12.1	15.6	22.4	11.5	17.7	26.6	12.4	18.6	25.2
35	18.0	23.4	31.7	12.6	15.3	23.0	12.3	18.3	27.1	12.8	19.3	25.6
30	20.0	26.4	35.8	14.0	17.1	25.8	13.9	21.1	30.8	14.1	20.6	28.3
25	22.4	28.6	40.1	15.5	19.7	29.0	15.5	22.9	34.8	15.8	21.5	31.1

Reallocation Scenario: focus on WhitefishN

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Present	15.6	20.7	28.3	11.1	14.2	21.8	10.7	16.6	24.3	11.5	17.2	23.6
45	17.0	25.7	35.4	12.5	16.2	21.2	12.1	18.2	29.2	11.5	16.1	20.6
40	18.2	27.1	37.4	13.4	17.2	22.4	13.1	19.6	31.1	12.0	16.9	21.8
35	19.4	29.5	41.0	14.2	18.8	24.5	14.0	20.9	34.2	12.6	17.7	23.5
30	21.6	32.8	44.0	15.7	20.5	26.4	15.5	23.2	37.1	13.9	20.3	25.4
25	'---	'---	'---	'---	'---	'---	'---	'---	'---	'---	'---	'---

3.4.2. Multi-Indicator compromise Neyman allocation

To circunvent the negative consequences of single-indicator optimization for the the remainder of indicators, multi-indicator compromise Neyman allocations were considered. The redistributions of sampling effort across strata obtained by this methodology are displayed in the following table with strata weights highligthed in parenthesis.

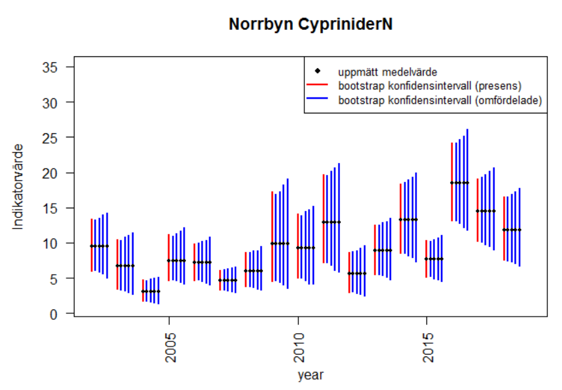
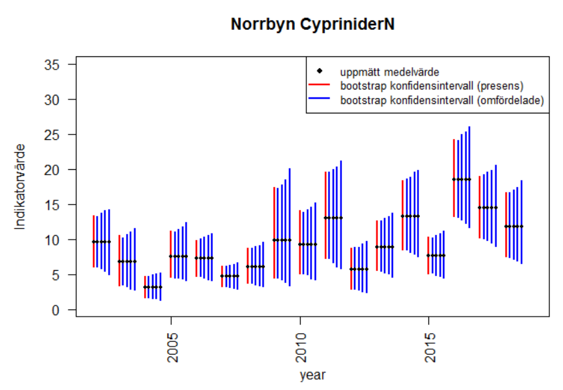
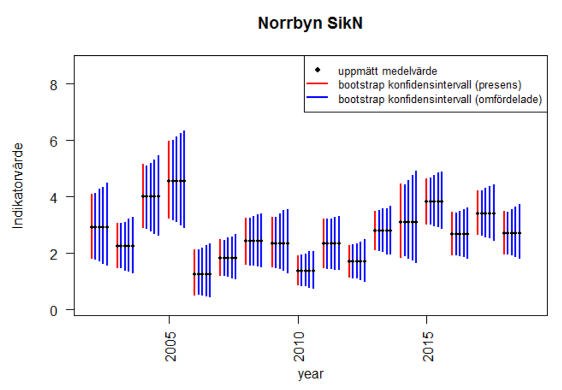
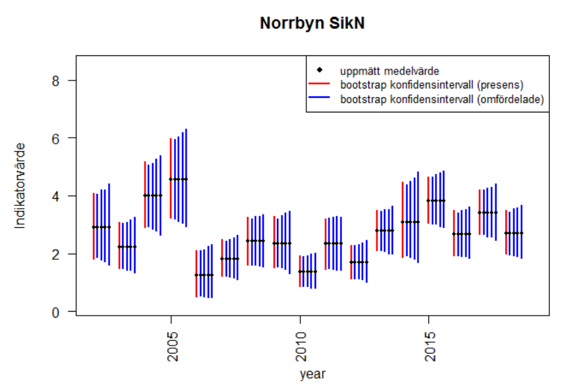
Depth strata	Present		Compromise0		Compromise1	
0-3 m	14	(0,31)	16	(0,35)	14	(0,31)
3-6 m	12	(0,26)	11	(0,24)	11	(0,24)
6-10 m	14	(0,31)	13	(0,39)	15	(0,33)
10-20 m	5	(0,11)	5	(0,11)	5	(0,11)

Where:

Compromise0 = AllocScenarioCypNHerNPerN
Compromise1 = AllocScenarioCypNHerNPerNWhiN

It is worth noticing that the weights of both compromise allocations were relatively close to the allocation scheme presently being used when sampling this area.

The following graphs display the evolution of the simulated confidence intervals of two contrasting indicators (rows) under the two compromise solutions (columns). Each graph displays the confidence interval of the original series (red line) and confidence intervals obtained with successively smaller sample sizes (blue lines). Full results for all scenarios and indicators are displayed in the table that follows. In this table values are expressed in terms of relative standard error (RSE) as calculated from bootstrap. To facilitate interpretation a colour code is used in the cells – Yellow when values are lower than those presently obtained (first row); Green when RSE are within +5% of present values; and red when RSE values are beyond that 5% of the present value.

	Reallocation Scenario	
	Compromise0 **[CypN + HerN + PerN]**	**Compromise1** **[CypN + HerN + PerN + WhiN]**
CyprinidерN		
SikN		

The graphs and table show that a reallocation of effort is **likely to improve the precision of results of some indicators but not much. Under such compromise2, a sample size reduction to 40 or even 35 would still largely provide approximately the same results as obtained with the present sample allocation scheme and sample size**. Under such effort reallocation and sample size reduction scenario, detrimental effects in the precision of the main indicators would have been minor relative to the values originally obtained.

Compromise0 **[CypN + HerN + PerN]**

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Present	15.6	20.4	28.1	11.2	14.1	21.8	10.8	16.5	24.3	11.3	17.2	23.6
45	15.7	20.3	27.6	11.1	14.1	21.9	10.8	16.3	24.3	11.7	17.7	23.7
40	16.9	21.5	30.0	12.0	14.6	22.2	11.7	17.8	25.4	12.5	18.7	25.7
35	18.1	22.6	32.1	12.8	15.9	23.9	13.1	18.9	27.1	13.5	20.1	27.4
30	20.2	25.6	35.1	14.1	16.9	25.9	14.1	21.1	30.0	14.5	20.9	29.4
25	22.3	28.7	39.6	15.6	18.9	29.1	15.8	23.2	33.8	16.1	22.1	32.9

Compromise1 **[CypN + HerN + PerN + WhiN]**

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
Present	15.7	20.5	28.3	11.2	14.1	21.7	10.8	16.5	24.3	11.3	17.2	23.6
45	15.8	20.9	28.6	11.1	14.2	21.4	10.6	16.3	24.6	11.4	17.3	22.7
40	16.9	22.2	30.1	12.0	15.0	21.9	11.4	17.6	25.8	12.3	18.0	24.2
35	18.3	23.3	32.6	12.8	15.8	24.0	12.6	19.0	27.7	12.9	19.4	26.0
30	20.1	26.3	35.6	13.9	17.1	25.7	13.9	20.9	30.7	14.1	20.7	28.4
25	22.2	29.3	39.7	15.3	19.4	28.6	15.3	23.1	34.6	15.8	21.3	31.2

It is worth noticing that the results obtained at area-level under the present sample size of 45 stations, a sample size of 39 stations with the present allocation (section 3.3) and a sample of 40 stations with the allocation suggested under compromise0 or compromise1 (this section) remained relatively similar (within a few percent points). These results are displayed in the following table where the most precise allocation for each indicator is highlighted in yellow.

	CyprinidsN			HerringN			PerchN			WhitefishN		
	5%	Median	95%	5%	Median	95%	5%	Median	95%	5%	Median	95%
45 (present)	**15.7**	**20.4**	**27.9**	**11.1**	**13.9**	**21.8**	**10.8**	**16.5**	**24.4**	**11.4**	**17.3**	**23.5**
39 (original)	17.2	22.1	30.5	**12.0**	15.3	22.2	11.6	**17.5**	26.4	12.4	18.7	24.7
40 (comp0)	**16.9**	**21.5**	**30.0**	**12.0**	**14.6**	22.2	11.7	17.8	**25.4**	12.5	18.7	25.7
40 (comp1)	**16.9**	22.2	30.1	**12.0**	15.0	21.9	**11.4**	17.6	25.8	**12.3**	**18.0**	**24.2**

3.5. Variability of trends with different allocation and sample size

The following slopes and results of slope significance test (H_0: slope=0, $p<0.05$) were determined for the present estimates at area-level

Indicator	Slope	Significance?
CyprinidsN	0,498777	TRUE
HerringN	0,368376	FALSE
PerchN	-0,97714	TRUE
WhitefishN	0,009242	FALSE

The next tables show the the number of replicates (out of 1000) that registered slope with same sign and the same outcome of slope significance test (as originally determined from present estimates) for different varying sample size and allocations. **In agreement with previous analysis, it is noticeable that if the samples had been re-allocated and the sample size reduced to 39-40 stations the general perception of the trends in the main indicators would not have differed.**

Original allocation

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN

Present	832	884	1000	627
45	846	861	1000	643
39	795	877	1000	630
34	746	872	998	619
31	711	868	997	640
25	639	826	995	604

Compromise0 **[CypN + HerN + PerN]**

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	832	869	1000	675
45	849	881	1000	672
40	812	875	1000	649
35	774	878	998	627
30	677	881	997	619
25	657	866	998	600

Compromise1 **[CypN + HerN + PerN + WhiN]**

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	863	890	1000	662
45	854	880	1000	672
40	804	855	1000	669
35	777	847	1000	621
30	690	848	999	624
25	632	862	997	638

3.6. Variability of trends with different allocation and sample size **(sampling every second year)**

The following tables show similar results when a change in sampling periodicity from annual to once every two years is simulated. It is clear the different results that would have been obtained, **particularly for CyprinidsN and PerchN.**

Original allocation

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	162	962	515	546
45	184	970	482	558
39	174	961	501	529
34	146	943	504	547
31	175	958	508	516
25	168	948	492	544

Compromise0 **[CypN + HerN + PerN]**

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	169	962	506	565
45	196	974	478	575
40	193	964	516	556
35	194	947	516	543
30	168	958	513	553
25	175	941	482	527

Compromise1 **[CypN + HerN + PerN + WhiN]**

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	161	959	509	551
45	171	976	482	570
40	167	962	520	551
35	200	956	518	541
30	179	948	496	536
25	190	943	471	549

3.7. Variability of trends with different allocation and sample size **(sampling every third year)**

The following tables show similar results when a change in sampling periodicity from annual to once every three years is simulated. It is clear the different results that would have been obtained, **particularly for CyprinidsN and PerchN but now also for HerringN and WhitefishN.**

Original allocation

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	223	705	94	416
45	217	702	90	438
39	218	724	93	448
34	195	736	101	440
31	216	718	117	460
25	207	725	109	471

Compromise0 **[CypN + HerN + PerN]**

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	209	708	105	411
45	213	711	92	441
40	228	708	92	452
35	212	725	87	437
30	215	694	115	453
25	166	733	109	462

Compromise1 **[CypN + HerN + PerN + WhiN]**

SampSize	CyprinidsN	HerringN	PerchN	WhitefishN
Present	220	715	96	409
45	229	702	92	422
40	226	707	91	439
35	214	719	103	443
30	211	716	101	459
25	191	708	117	449