Survey effort allocation using advanced design: clam population as case study

Nathalie CAILL-MILLY[a], Noëlle BRU[b], Frank D'AMICO[b], Florence SANCHEZ[a], Muriel LISSARDY[a] , Jennifer BROWN[c]

[a] Ifremer, FED 4155 – MIRA, Laboratoire Ressources Halieutiques Aquitaine, France
[b] UNIV PAU & PAYS ADOUR, Laboratoire de Mathématiques et de leurs Applications de Pau, UMR CNRS 5142, FED 4155 MIRA, France
[c] University of Canterbury, Department of Mathematics and Statistics, Christchurch, New Zealand

Marine monitoring

3rd International conference in Aarhus, Denmark, 1-2 October 2015

- Management and long-term conservation of a natural resource need to take into account its spatial distribution
- For such purpose, designing an efficient and robust monitoring program is an essential step
- For large areas

Probability-based survey designs		Spatially-balanced survey designs
• Random Stratified Sampling (StRS) • Systematic sampling • ….. • Commonly used	→	• Generalized Random Tessellation Stratified sampling (GRTS) • Balanced Adaptive Sampling (BAS) • …. • Becoming more popular

Main differences….

Stratified random sampling (StRS)

➢ Provides some spatial structure to the overall population through strata

➢ Each stratum is sampled independently, commonly using Simple random sampling (SRS)

➢ SRS is used to sample from a population *S* by generating a series of random locations, *x,y* values, that are paired to form a set of *s* random sample locations

➢ Advantages: simple and flexible, additional samples can be easily added to an existing set of samples

➢ Disadvantages: in some cases, existence of clusters of samples or areas devoid of samples

Generalized Random Tessellation Stratified sampling (GRTS)

➢ View also as a stratified survey design, it provides a spatially-balanced design

➢ GRTS method involves different steps:

- Creating a grid with hierarchically ordered addresses
- Randomizing them
- Finally sampling them with a reverse hierarchical ordering function (quadrant-recursive ordering using Morton order)

➢ Developed for monitoring natural resource trends since the year 2000. For aquatic resources, mainly applied for discrete points and linear networks (rivers…)….

From Stehman, 1999; Stevens and Olsens, 2004; Herlihy et al. 2000; Theobald et al., submitted; Saalfeld, 1998

Main differences….

Stratified random sampling (StRS)

➢ Provides some spatial structure to the overall population through strata

➢ Each stratum is sampled independently, commonly using Simple random sampling (SRS)

➢ SRS is used to sample from a population *S* by generating a series of random locations, *x,y* values, that are paired to form a set of *s* random sample locations

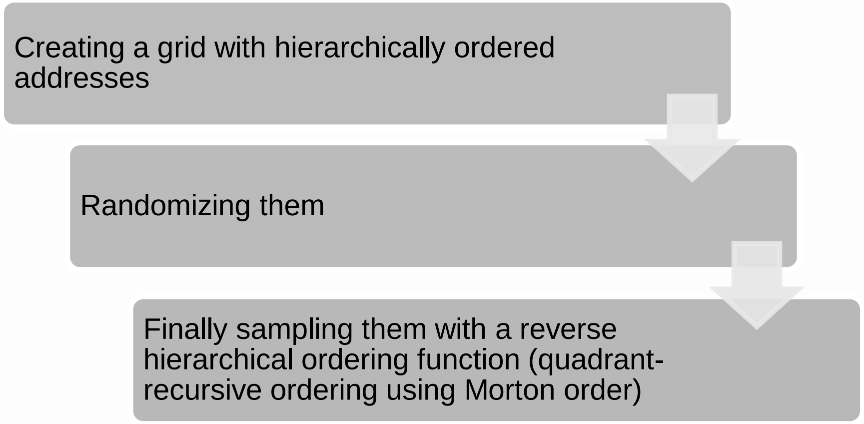
➢ Advantages: simple and flexible, additional samples can be easily added to an existing set of samples

➢ Disadvantages: in some cases, existence of clusters of samples or areas

Generalized Random Tessellation Stratified sampling (GRTS)

➢ View also as a stratified survey design, it provides a spatially-balanced design

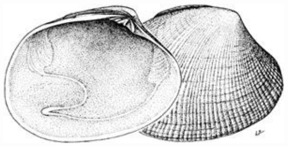
➢ GRTS method involves different steps:



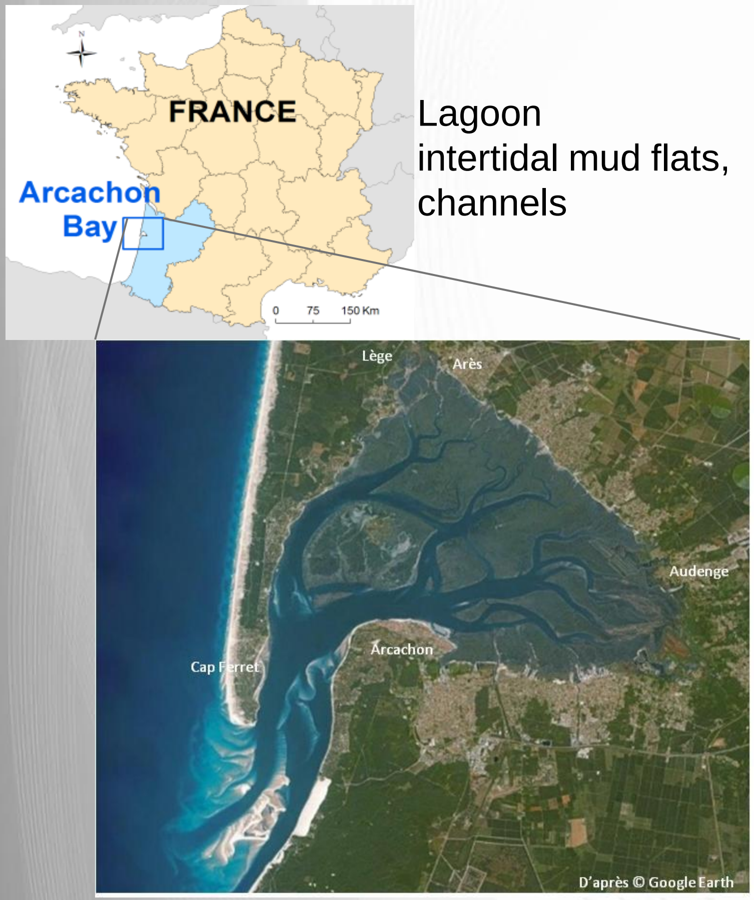
➢ Developed for monitoring natural resource trends since the year 2000. For aquatic resources, mainly applied for discrete points and linear networks (rivers…)….

=> The aim of this presentation is to compare StRS and GRTS used for a resource distributed in patches in area polygons
Case study: exploited bivalve population monitoring in an intertidal lagoon

In the case of Manila clam (*Venerupis philippinarum*) from Arcachon Bay



Source : FAO



Lagoon
intertidal mud flats,
channels

- Fishermen go on fishing areas by boat.
- Exploitation takes place at low tide, by hand (mixing of the sediment). They sometimes use a small tool to scrape the sediment (in the presence of old oyster shells for example).

EU legislation

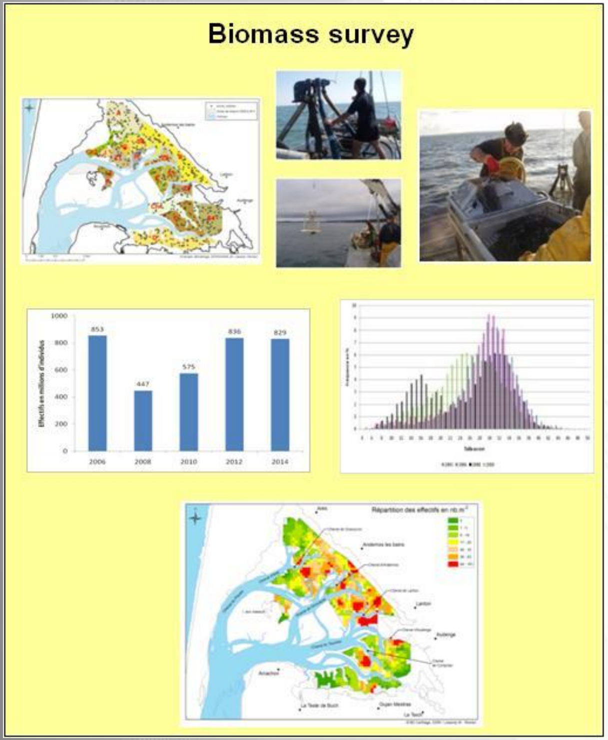
Minimum harvest size
35 mm since 2008

Local legislation

Protected areas

Limited number of licenses

Possibly limited fishing periods



To identify important issues for management decisions, a long term monitoring program exists with a close partnership between scientists and fishermen

Among the developed tools, a dedicated standardized survey campaign is undertaken since 2003 at the bay scale.

EU legislation

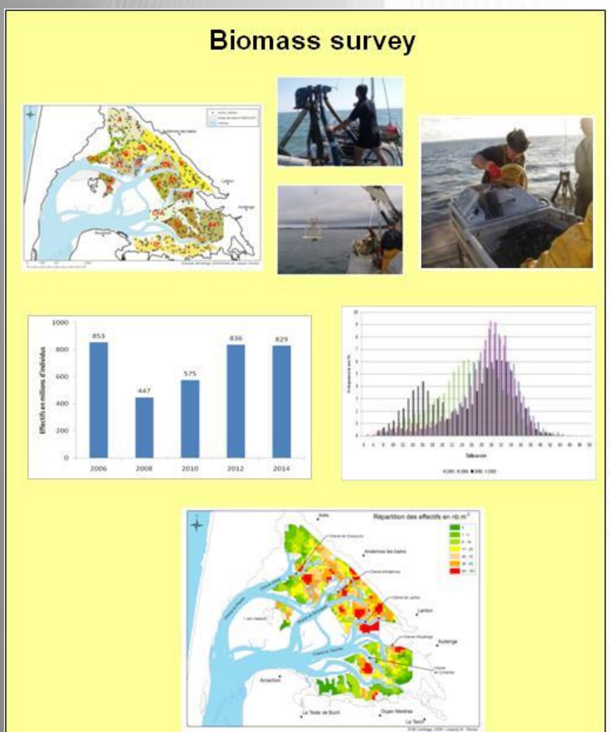
Minimum harvest size
35 mm since 2008

Local legislation

Protected areas

Limited number of licenses

Possibly limited fishing periods



But,
- time consuming;
- relatively costly.

⇒ Logical to consider way to be **more efficient (reduction of the number of stations without loss of precision)**

⇒ One way is to look at our sampling protocol

Step 1	**Known population: abundance and spatial distribution**
Step 2	**For both designs, selection of samples: locations and optimal number of stations**
Step 3	**Comparison of performances**

Step 1

Biomass and density data 2012 obtained using StRS protocol → *Kriging* / *RGeos* → Virtual population

Division of the area into strata

+

Step 2

Site selection through 2 probabilistic sampling design

SRS	*GRTS*
sp	*spsurvey*
Site selection (for each stratum)	Site selection (for each stratum)
Kriging grid — *RGeos*	*RGeos* — *Kriging grid*
Estimation of Biomass or Density for each selected site (within each stratum)	Estimation of Biomass or Density for each selected site (within each stratum)
Berthou el al. equations	*Berthou el al. equations*
Estimator performance : variance	Estimator performance : variance

Step 3

Segmentation on performance curve — *segmented*	*segmented* — *Segmentation on performance curve*

Comparison of sampling design performance through optimal number of selected sites given estimated parameter (biomass vs density)

In blue, used R packages and functions

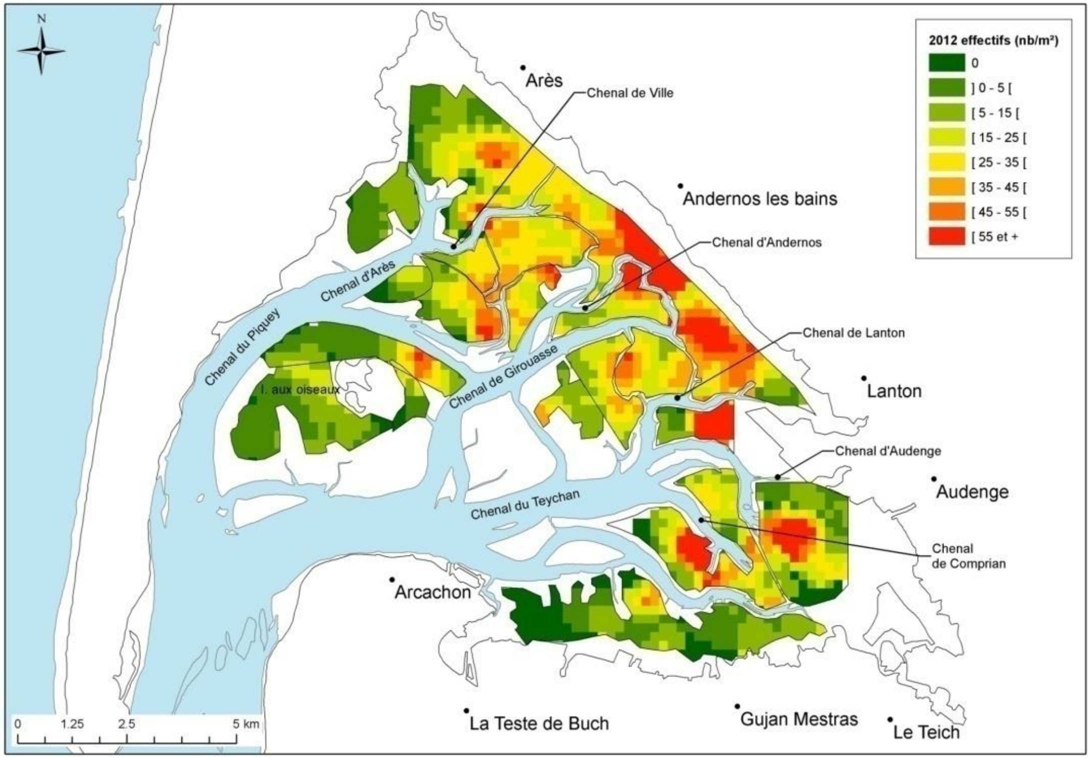
Step 1

Biomass and density data 2012 obtained using StRS protocol → Virtual population → Division of the area into strata

Using geostatistical analysis (ordinary kriging), construction of a virtual population

The area is assumed to be a grid. For each cell, one every 200 m, an estimated value is assigned. Values are expressed in number (abundance) and in gramm (biomass) per m²

Used package/function: Rgeos



Density expressed as abundance per m²

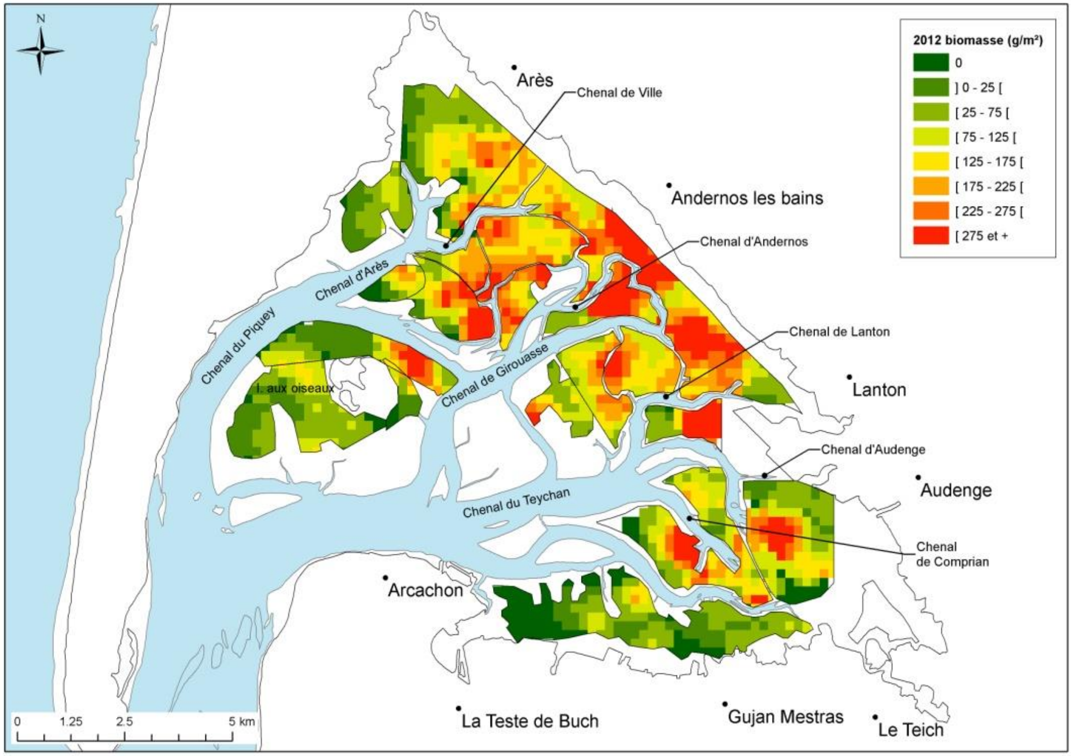
Step 1

Biomass and density data 2012 obtained using StRS protocol → Virtual population → Division of the area into strata

Using geostatistical analysis (ordinary kriging), construction of a virtual population

The area is assumed to be a grid. For each cell, one every 200 m, an estimated value is assigned. Values are expressed in number (abundance) and in gramm (biomass) per m²

Used package/function:
Rgeos



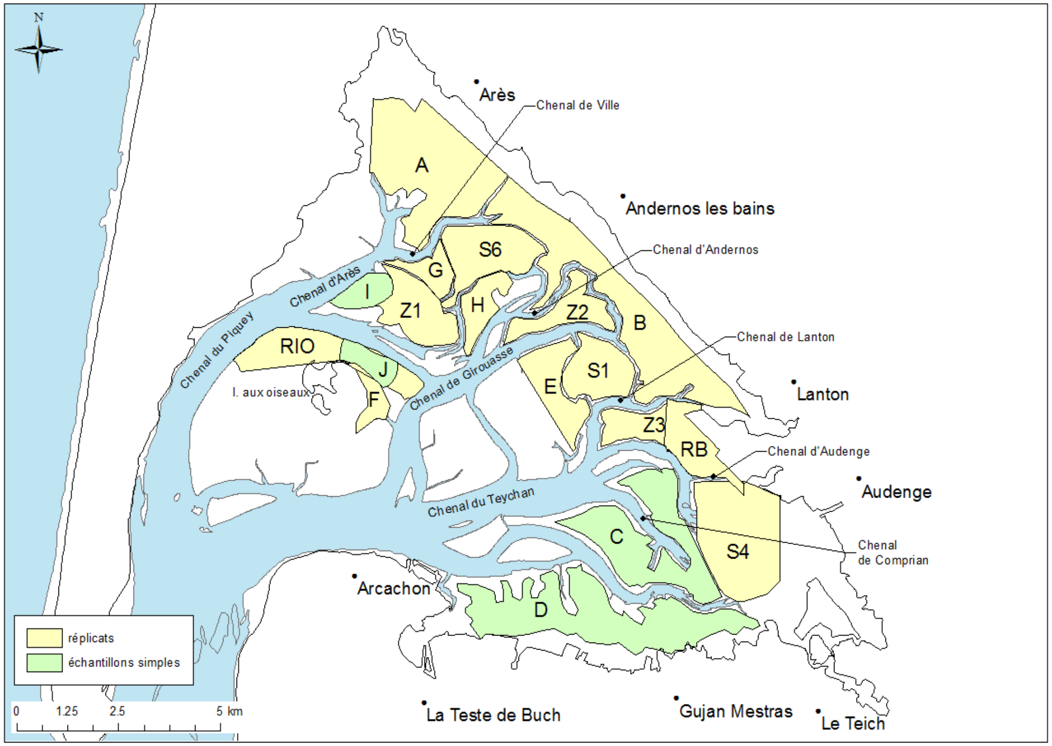
Density expressed as biomass per m²

Step 1

Biomass and density data 2012 obtained using RSS protocole → **Virtual population** → **Division of the area into strata**

A spatial division is applied according to environmental and regulation conditions
⇒ 18 strata are defined

Each stratum has to be sampled for management purposes



Spatial division with strata

Step 2

Site selection through 2 probabilistic sampling design

SRS	***GRTS***
↓	↓
Site selection (for each stratum)	**Site selection (for each stratum)**
↓	↓
Estimation of Biomass or Density for each selected site (within each stratum)	**Estimation of Biomass or Density for each selected site (within each stratum)**
↓	↓
Estimator performance : variance	**Estimator performance : variance**

For each stratum, for the two types of allocation effort samples:

- Increasing the sampling effort from 1 to n_i (i representing a given stratum), run of 1000 simulations

- At each step, estimation of the densities (abundance and biomass). Precision of estimators is computed using variance formulae

To compare the efficiency of the 2 methods: decreasing plot of the variance vs number of sampling sites

Used package/function:
For SRS, spsample of the R sp package
For GRTS, grts of the R spsurvey package

Step 3

Comparison of sampling design performance through optimal number of selected sites given estimated parameter (biomass vs density)

For SRS and GRTS, the optimum number of sampling stations is considered to be achieved following this process:

- Fit regression models with segmented relationships between the response (variance) and the explanatory variable (number of sites);

- So estimated break-points and slopes are identified;

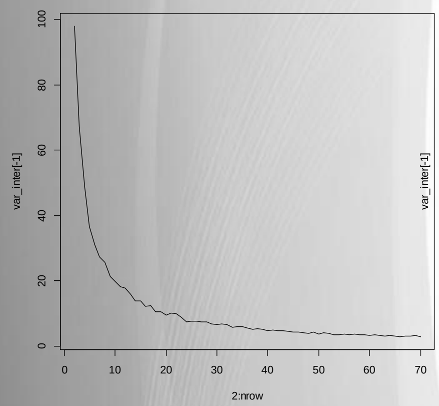
- Then, as soon as the slope of a given segment does not differ from zero, the corresponding breakpoint is assumed to be the optimal number of sampling stations.

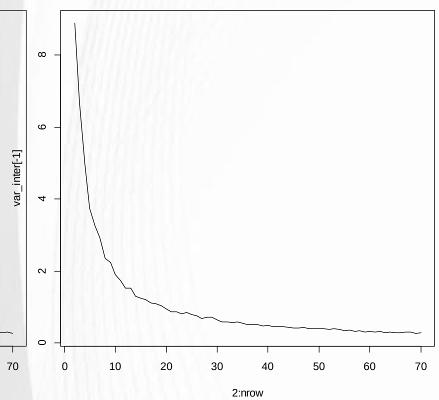
Used package/function:
change.point and slope of R segmented package

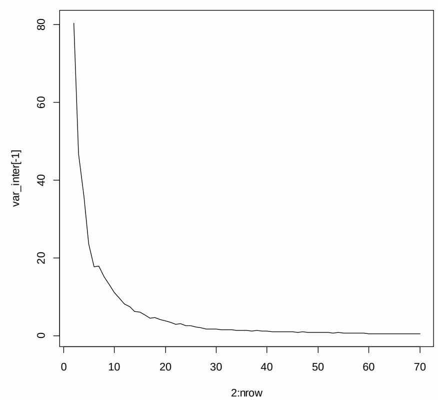
SRS

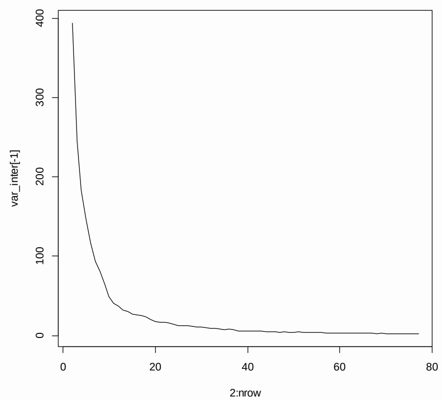
GRTS

Stratum A ***Stratum B*** ***Stratum A*** ***Stratum B***

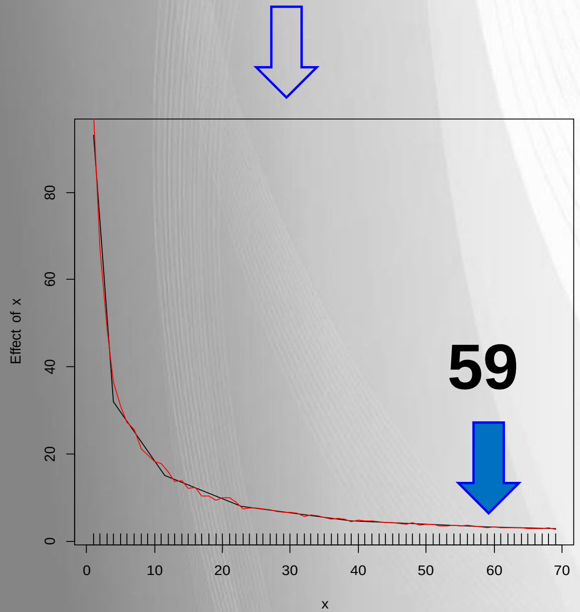


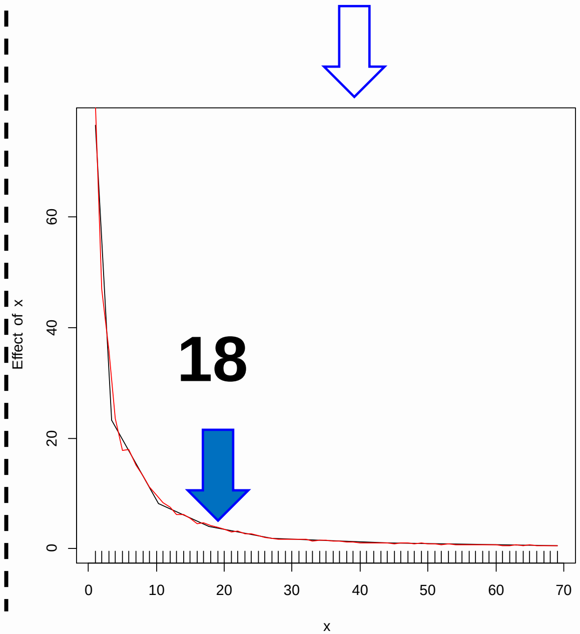






Examples of plots of variance vs number of sampling sites - Abundance





For each stratum, identification of the optimum number of stations with the associated variance
Note that n_i is the previous maximum number of stations used during the scientific surveys

SRS

			Biomass (g.m²)			Abundance (nb.m²)		
Stratum	Surface km²	n_i	n optim	SE (n optim)	*Precision*	n optim	SE (n optim)	*Precision*
Z2	1,59	17				15	40	*58*
E	1,96	21						
S1	2,00	21	17	5	*326*			
Z1	2,20	25	17	7	*1033*			
S6	2,69	29	24	6	*113*	25	7	*3*
RIO	3,19	34				30	5	*8*
C	4,89	49	47	15	*134*	32	3	*9*
S4	4,98	51	32	8	*474*			
A	6,71	70				59	22	*3*
D	6,90	70	53	23	*6*	61	11	*1*
B	7,33	77				55	15	*16*

GRTS

			Biomass (g.m²)			Abundance (nb.m²)		
Stratum	Surface km²	n_i	n optim	SE (n optim)	*Precision*	n optim	SE (n optim)	*Precision*
Z2	1,59	17	11	~ 0	*492*	11	1,	*23*
E	1,96	21	7	1	*558*	6	~ 0	*12*
S1	2,00	21						
Z1	2,20	25	16	3	*380*	17	10	*5*
S6	2,69	29	20	1	*57*	20	4	*2*
RIO	3,19	34	23	3	*134*	6	1	*21*
C	4,89	49	19	4	*139*	14	2	*11*
S4	4,98	51	33	12	*93*	28	4	*5*
A	6,71	70	16	3	*158*	18	4	*5*
D	6,90	70	23	1	*6*	23	3	*1*
B	7,33	77	18	4	*130*	19	4	*20*

=> Whatever the protocol, this simulation process shows that the number of stations per stratum can be reduced without loss of precision

=> It is not possible to assess an optimal number of stations for small sized-strata (i.e. < 2 km², results not shown) excepted for 2 strata (Z2 and E)

=> Biomass estimation vs Abundance estimation bring the same conclusion
GRTS generally leads to a smaller optimum number of stations (except mismatch for strata RIO and E)

Take home messages

Compared to ACTUAL sampling effort, GRTS should provide greater performance than SRS whatever biomass or density is assessed

Sampling effort on small strata has to be reconsidered

Nowaday
15 days at sea
1 scientist, 2 fishermen, 3-4 crew member
~ 50 k€ per campaign (373 000 DKK)

→

Future
10 days at sea
1 scientist, 2 fishermen, 3-4 crew member
~ 35 k€ per campaign (261 000 DKK)

Short terms

Check if those results are confirmed when using others campaigns' data

Refine the benefit for the campaign's organization

Organize meeting with stakeholders (fishermen, administration) to present the tested method and convince them about its soundness
Mandatory to go-on the co-management approach we advocate

Medium terms

Go further and test other recent spatial balanced sampling designs (e.g. BAS)

Consider if same results are obtained with bivalves presenting slightly different distribution characteristics (size of the patches…), e.g. mussels, cockles….

For any detail on statistics, my colleague, Noëlle Bru, will help me to answer.

She is a field staticician !!!

Thanks for your attention