Presentation given at ICES WG BIFS 2005

SOME PROBLEMS OF INTERNATIONAL SURVEYS IN THE BALTIC SEA

Кasatkina S.M., Gasyukov P.S.

Atlantic Scientific Research Institute of Marine Fisheries and Oceanography (AtlantNIRO). 5, Dm.Donskoy Str., Kaliningrad, 236000 Russia .Fax: 007 0112 21999,

e-mail: ks@atlant.baltnet.ru

With support of BSRP LL on ICES Joint Open Sea Surveys

BACKGROUND

The survey results provide major information for:

- ***studying the spatial and temporal variability of abundance and biomass distribution***
- ***assessment of abundance indices applied in tuning integrated statistic methods of commercial fish stocks assessment (ADAPT, Gavaris, 1988), (XSA, Shepherd, 1991,1999), (ICA, Patterson, 1994).***

Uncertainty of estimates are not less important than the abundance and biomass estimates, since the decisions on fishery management are adopted on their basis.

Quantifying, summarizing and integrating total uncertainty in fisheries resource surveys" is one of the major terms of references of ICES Working group (ICES FTC Report, 2003, 2004; WGFAST, 2004; WGSAD ,2004).

Our aims are addressed to following aspects:

- how abundance uncertainty estimates could be applied in the current methods of stock assessment;

- methodical aspects of acoustic sampling uncertainty assessment on the basis of simulation

- the problems which, in our opinion, are of primarily importance for the acoustic surveys improvement in the Baltic Sea

Surveys data as an important component of advanced stock assessment methods

The main method of commercial stocks assessment in the Baltic Sea:

The the extended survival analysis (XSA) developed by Shepherd (1991, 1999) and belonged to the class of methods structured by age groups

The important component of the input data:

Abundance indices by age and fishing years describing the temporal dynamics of commercial fish population abundance

More preferable sources of these abundance indices:

The long-term surveys as a results of scientific measurements

The unknown parameters of the model:

The abundances of survived fish at the end of the terminal year

The objective function

$$SSE\left(S_{1,y_k},\ldots,S_{a_k,y_k},S_{a_k 1},\ldots,S_{a_k,y_k-1}\right)=$$

$$\sum_{f=1}^{n_f}\sum_{a=1}^{a_k}\sum_{y=1}^{y_k}\left\{\ln\left(P_{a,y,f}\right)-\ln\left(N_{a,y,f}\right)\right\}^2/\sigma^2_{a,y,f}$$

$S_{a,y}$ -**number of survived fish at the age at the end of the year ;**

$\sigma^2_{a,y,f}$ -**variance of logarithm of abundance index of the age group in the year *y* for the index *f* ,**

$$P_{a,y,f}=\frac{1}{q_{a,f}}\cdot I'_{a,y,f}$$

$I'_{a,y,f}$ - **abundance index *f* for the age group *a* in the year *y*, recalculated by the beginning of the year.**

On the one side

The variance value, being the function of age, year and abundance index number, appeared in the equation denominator under the summation symbol.

The abundance indices standardization using generalized linear models requires *a priory* knowledge of the probable relationship between the variance and mean value (McGullagh, P., Nelder, J.A. 1989).

On the other side

Currently, objective function realized in ICES program (Darby, Flatman, 1994) which was used to assess the stock of all commercial fishes in the Baltic Sea is based on :

the assumption that the abundance index variance is a constant value for each the age group and each abundance index.

Let us demonstrate the relationship between the abundance index variance and abundance index value. The bottom trawl survey data in the Baltic Sea will be used.

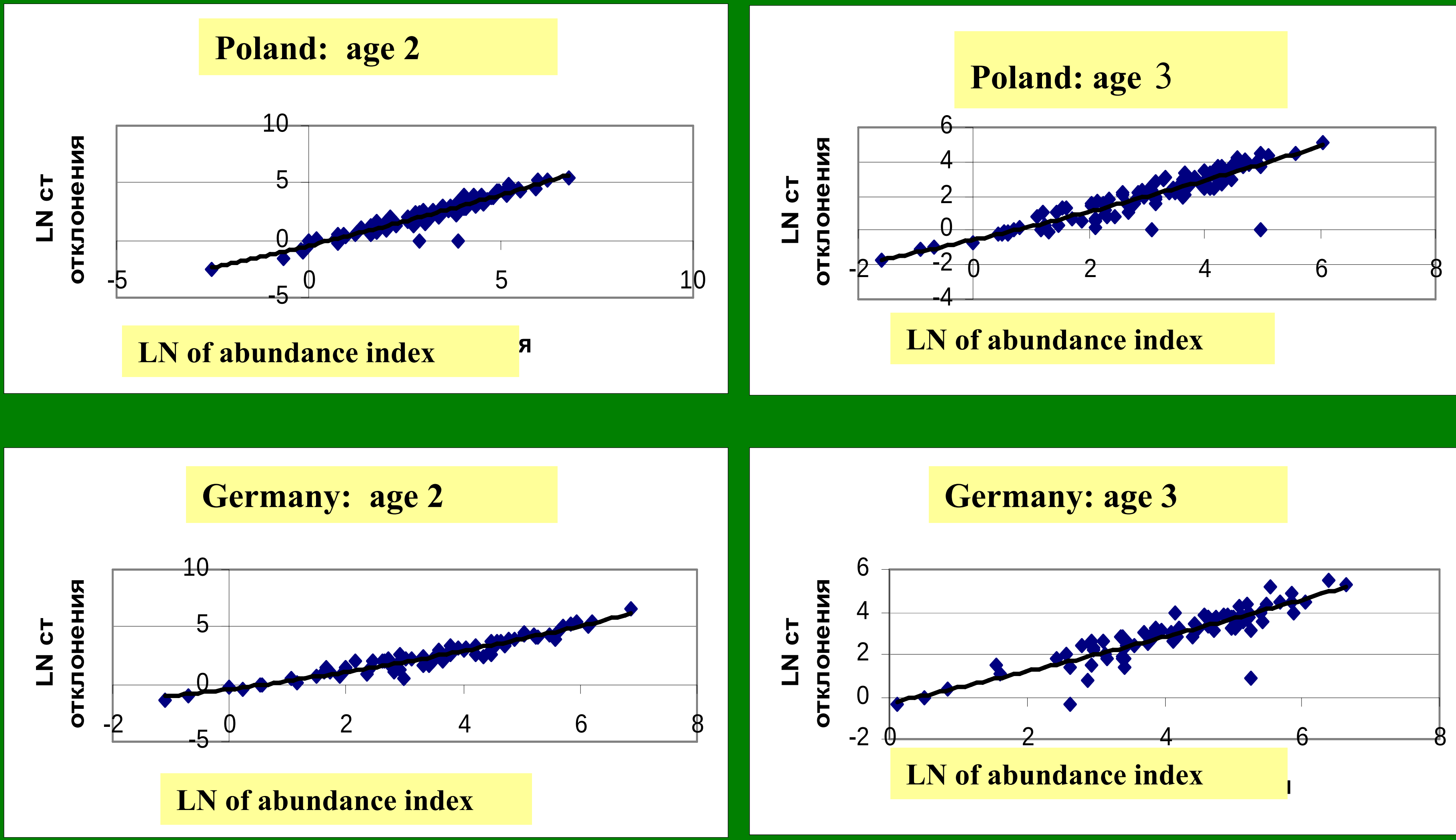
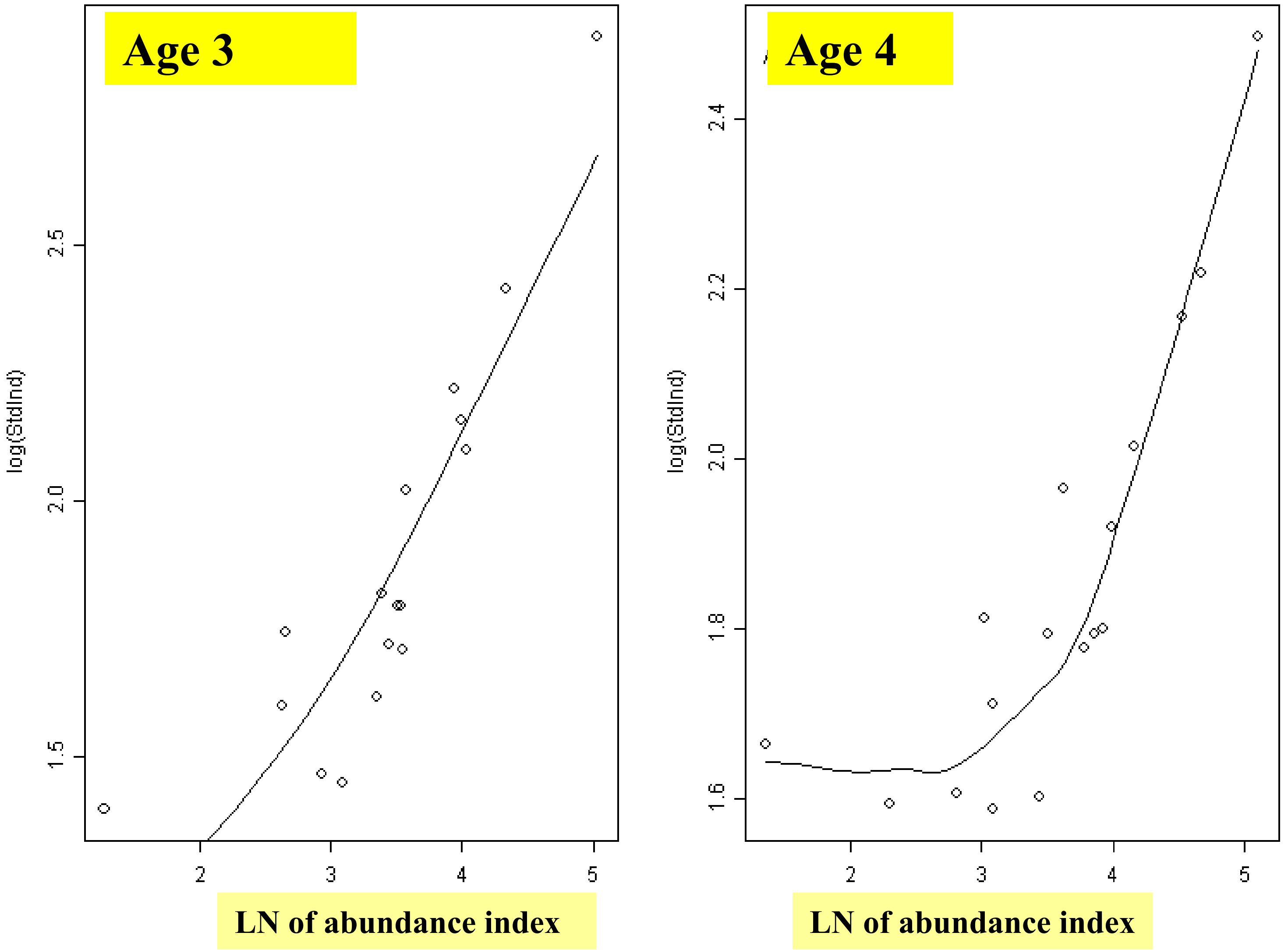


Figure 1. Dependence of logarithm of abundance index standard deviation of cod abundance as a function of abundance index logarithm estimated based on series of stations fulfilled by Germany (1980-1999, subdivisions 25, 26, 28) and Poland (1981-1999, subdivisions 25, 26)



Data from

Denmark, Sweden,

Germany,

Poland,

Russia and Latvia in 1982-1999.

Fig 2. Relationship between logarithm of the abundance index standard deviation and logarithm of the abundance index for standardized abundance index (GLM model based on the international trawl surveys results)

The results presented prove that the abundance index variance depends on abundance index value, hence variance could not be constant from year to year for each age-group.

Apparently , the stock assessment model accounting for this relationship, is able to describe the stock dynamics of the main commercial fishes in the Baltic Sea more precisely.

Realization of this model for cod in subdivisions 25-32 is possible already now, and it seems to be implemented in the nearest future.

What is about acoustic survey?

The stock assessment model accounting for variance relationship is not possible for the pelagic fishes of the Baltic Sea in view of non-availability of abundance indices accuracy estimates.

However

The relationship between the abundance index variance and abundance index is inherent not only in the estimates obtained on the basis of the bottom trawl survey data.

The same relationship is nature for the estimates based on the acoustic survey data, since it reflects spatial distribution of various fish species (Rivoirad et al, 2003; Simmonds et al, 2004)

Sufficient results of uncertainty estimates have been already obtained in fish assessment (Rivoirad et al,2000;Simmonds, 2004;Demer, 2004, Kasatkina and Gasyukov,2004; O'Driscoll, 2003).

ASSESSMENT OF UNCERTAINTY OF ACOUSTIC SURVEY RESULTS

Generally, overall survey uncertainty is estimated as overall variance, stipulated by:

- **sampling survey uncertainties and**
- **measurement survey uncertainties**

(*ICES WKSAD Report 2004; ICES WGFAST Report, 2004*)

MEASUREMENT SURVEY UNCERTAINTY **(or instrumental uncertainty)**

This is a collective term for all of the measurement components, associated with estimates of fish density at the each points along transect (ICES WG SAD Report, 2004):

- **instrumental error of acoustic equipment (calibration, weather conditions, bubble attenuation, signal threshold, fish delectability above the seabed and near the surface, and other),**

- **instrumental error of trawl tool (catchability and selectivity properties with respect to size-species identifications),**

- **fish behavior and avoidance reactions with respect to vessel and trawl, etc.**

At present neither of the works cited estimated measurement survey uncertainty, mainly due to the problems of accounting the sources of errors related to:

- **fish behavior**
- **the trawl as the sampling tool for size-species identifications**

In most cases, measurement uncertainty is been associated to the errors of calibration process and fish delectability .

In modern research echosounders the uncertainty associated with calibration and acoustic hardware is estimated to be ±0,2 dB (O'Driscoll, 2002), (Demer, 2000, 2004).

In most cases, the sampling survey variance is considerably exceeds the measurement survey variance (Rose et al, 2000; Demer, 2004)

In the modern acoustic surveys practice the uncertainty assessment in acoustic biomass estimates is mostly fulfilled as a <u>sampling survey variance</u>.

SAMPLING SURVEY UNCERTAINTY **is related to the errors caused by spatial variability** (WGFAST Report, 2004; WGSAD Report,2004):

- **acoustic index NASC (m2/m.mile2)**
- **species composition and length structure of fish species**

and estimated

by means of pooling the acoustic and biological data variance (Simmonds, 2004; Kasatkina and Gasyukov, 2004; Rose *et al*., 2000).

The pooling required development of *simulation methods*, where variability of different sources was assessed mainly with:

- **bootstrap resampling methods (Smith, 1997), (Robotham and Castillo, 1990)**
- **or geostatistical procedures (Petitgas, 1999), (Rivoirard et al, 2000).**

Below the simulation method developed at AtlantNIRO is considered, as well as a practical example of its application (Kasatkina and Gasyukov, 2004).

SIMULATION METHOD OF UNCERTAINTY ASSESSMENT IN ACOUSTIC SURVEY RESULTS

The method of sampling survey uncertainty assessment in biomass and abundance estimates has been developed on the base of:

- **acoustic survey data processing algorithms**
- **simulation with Monte Carlo method**

A major uncertainty sources :

- **spatial variability of acoustic index NASC (m2/m.mile2)**
- **spatial variability of species composition and length structure**
- **uncertainty of the target strength**

The effect of each uncertainty source was simulated applying the procedures of the bootstrap resampling method (Efron, Tibshirani, 1993).

Simulation of the replication set of each affecting factor is made in compliance with obtained empiric distribution functions for each stratum and subsequent processing by the stratified analysis

The proposed model allows to estimate the contribution of each uncertainty source and their cumulative effect (Kasatkina and Gasyukov, 2004).

The input data for simulation:

1. Regression equation of the target strength with statistic characteristics of parameters, including mean and standard deviation.

2. For each stratum
   - NASC values (m2/m.mile2) at all transects for each stratum
   - Length frequencies of all fish species, recalculated by the total catch, for each trawl station
   - Data used in fitting the "length-weight" relationship

3. Strata areas
4. Number of realizations in the simulation process

The inclusion of the target strength uncertainty into the system of affecting factors is the principle attribute of the proposed method of sampling uncertainty assessment

The latter is attained by means of the target strength regression equation accompanied by statistical characteristics of its parameters.

TS= a log L – b **mean a, sd (a)**

mean b, sd (b)

The target strength uncertainties are accounted :

- **from the variation in fish size**
- **from the error of regression parameters**

The practical realization of the developed method is illustrated below by the example of processing the data of icefish acoustic survey in the South Georgia area (R/V «*ATLANTIDA*», February 2002).

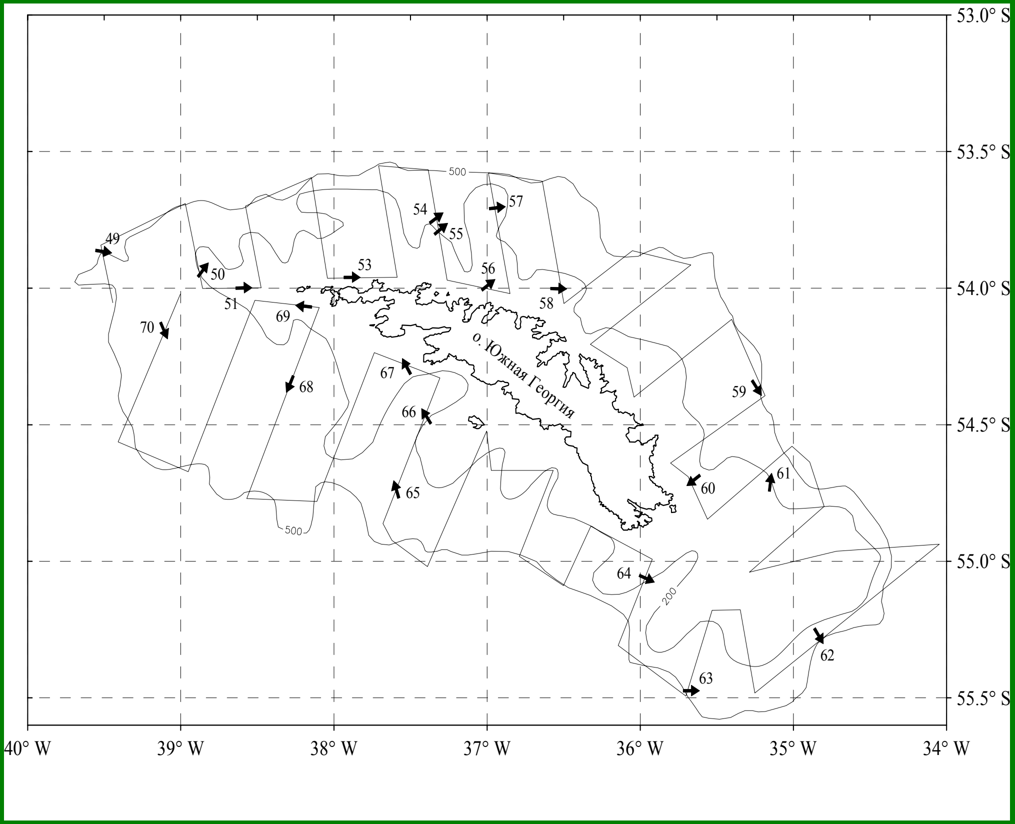
the South Georgia Islands

R/V «*ATLANTIDA*»,

Acoustic survey design

Stratified random sampling as randomly spaced parallel transects within each stratum.

The scheme of parallel transects within each stratum was based on two-step randomization (Kasatkina et al, 2002).



Acoustic sampling:

- Simrad EK-500 echosounder (38 kHz, 120 kHz)
- SonarData Echoview software

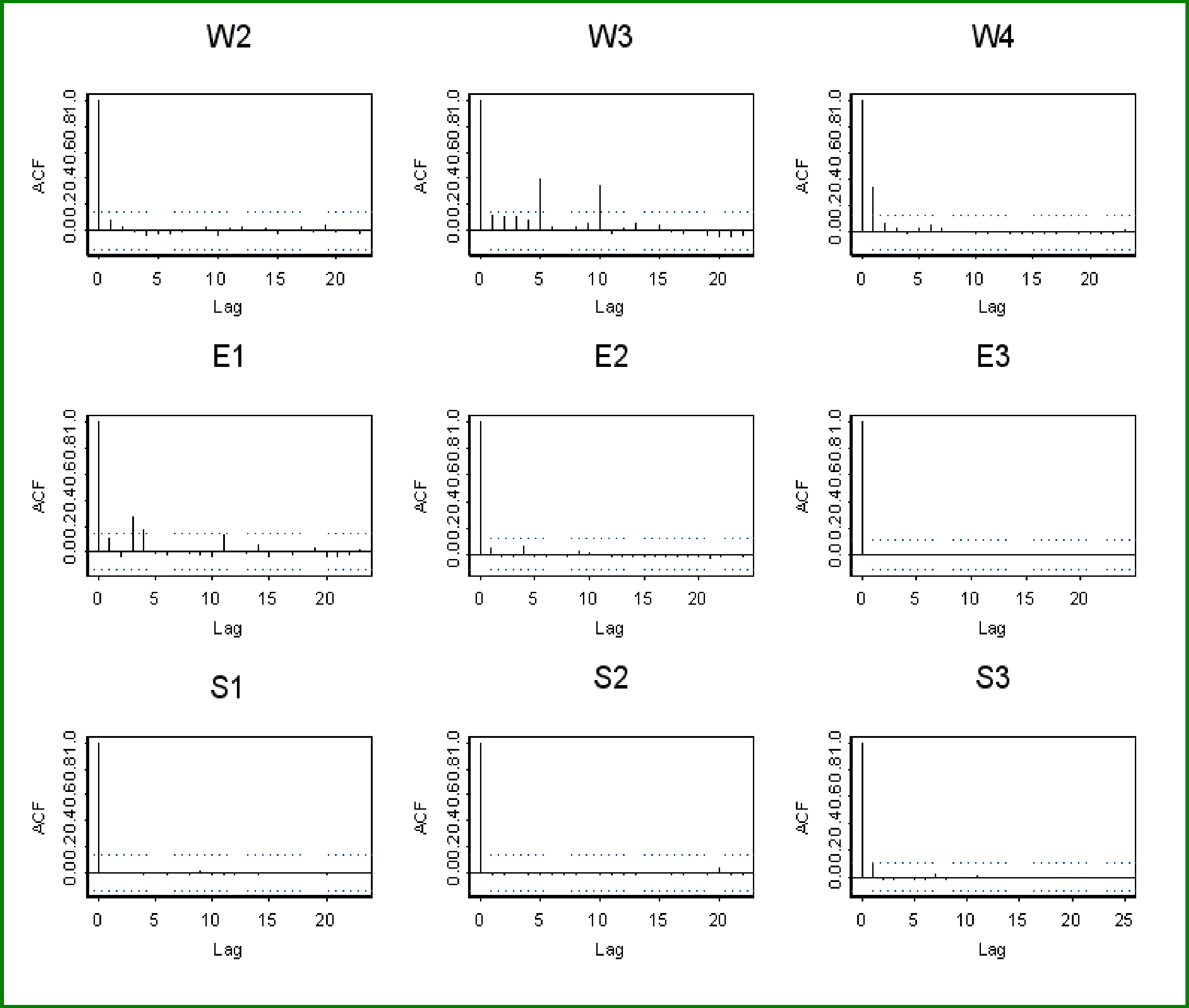
- Two frequency method for identification backscattering attributed to **fish fractions**

- Fish density were estimated on 38 kHz.

Trawl sampling

midwater trawl RT 70/308m

(vertical opening 43±3m)



Stratified random sampling in survey design allowed to exclude inter-transects correlation of acoustic samples.

Fig .3 presents autocorrelation functions of NASC values by the example of several transects in each stratum

The lack of autocorrelation allows to use *descriptive methods* in acoustic data processing.

Otherwise, if autocorrelation is encountered in acoustic data, *geostatistical analyses* would be more appropriate .

Target strength

TS in situ measurements :

- 38 kHz (EK-500, SonarData EchoViev software
- Only targets detected within 1° of the beam axis and within fished depth range .
- fish species amounted to at least 85% of the catch (in abundance).

Parameters of TS regression were estimated with the bootstrap method on the basis of 1000 replications

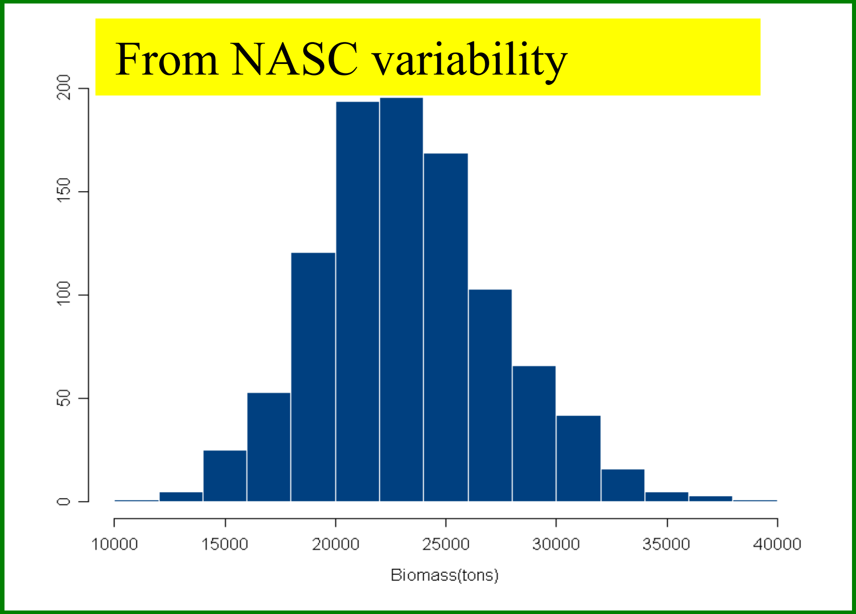
$$\mathbf{TS = 20 \log L + b_{20}}$$

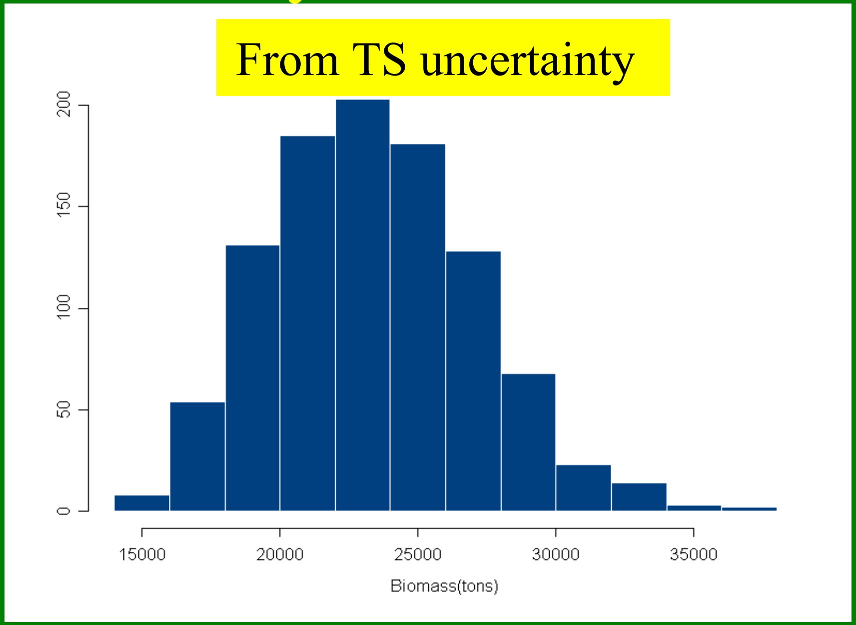
b_{20}=-83.61 dB, standard deviation sd=0.068 dB

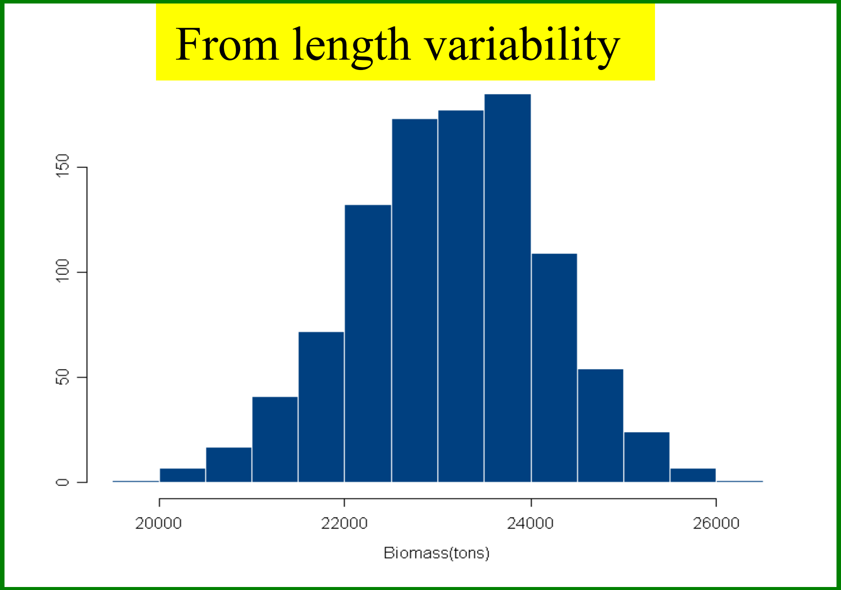
Statistical characteristics of fish biomass obtained by simulation

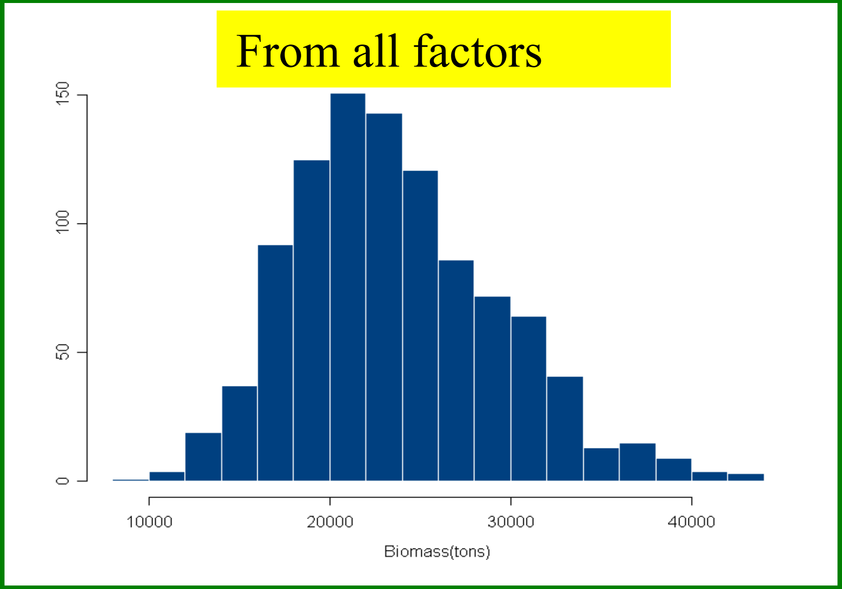
Parameters	Uncertainty sources			
	Uncertainty of NASC distribution	Target strength	Heterogeneity of fish length and species composition	All factors
Mean	**23659**	**27995**	**23131**	**27859**
Median	**22993**	**27999**	**23149**	**27727**
Standard deviation	**4122**	**4013**	**1040**	**5001**
Coefficient of variance	**0,177**	**0,014**	**0,045**	**0,180**
Lower limit	**15929**	**27215**	**21003**	**15063**
Upper limit	**31832**	**28799**	**25106**	**28012**
Lower limit of one-sided 95% confidential interval	**16872**	**27313**	**21350**	**2011**

Frequency distribution functions of biomass in relation to different uncertainty sources









Significance of statistical characteristics of TS regression parameters

Simulation results of biomass estimates from the same survey data by using TS regression with different uncertainty of coefficient b20

TS= 20 Log L – b20

TS= 20 log L – 83, 61 sd=0.068 dB (1)

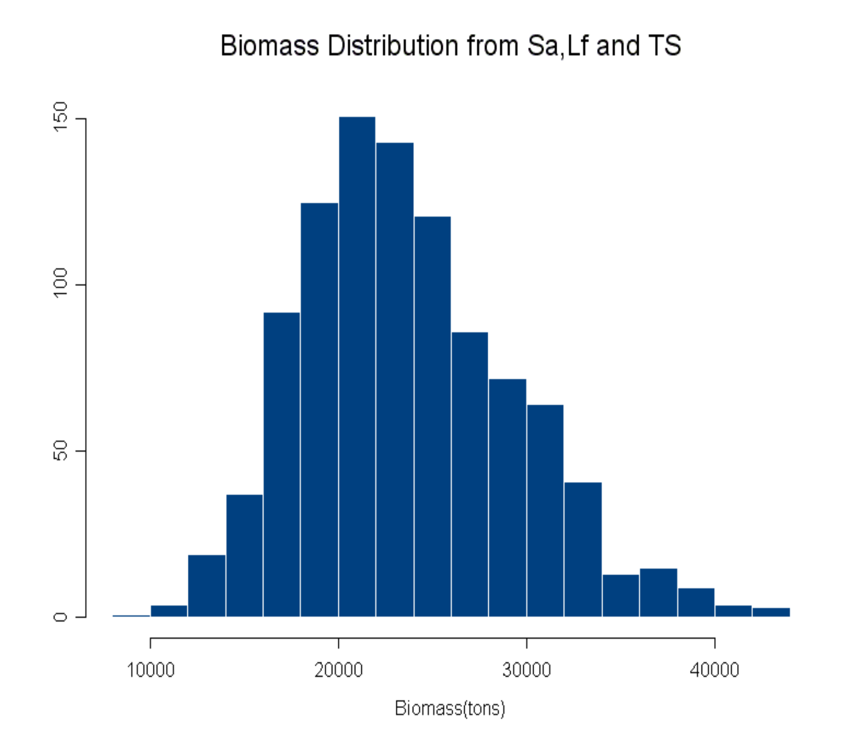
TS= 20 log L -83,41 sd=1.43 dB (II

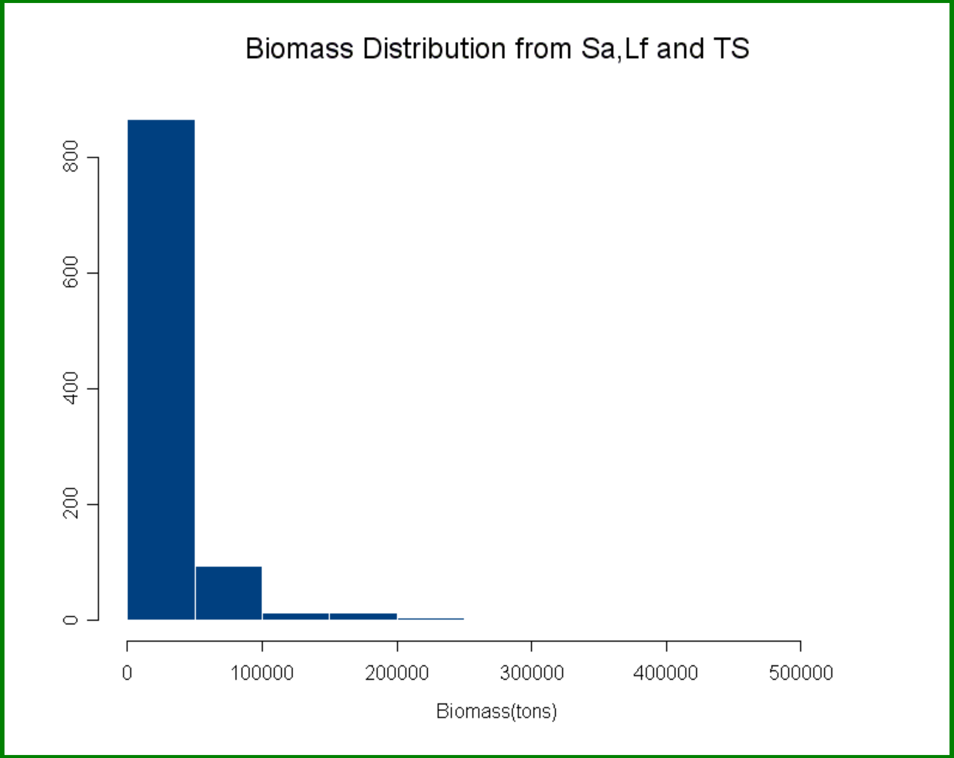
Two versions of biomass assessment

Parameters	TS regression (I) Sd= 0,068dB Combined	TS regression (II) Sd= 1,4 dB Combined
Mean	27859	35396
Median	27727	24487
SD	5001	41169
CV	0,180	1.163
Lower 95 % CL	15063	11532
Upper 95 % CL	28012	166738
OneSided95LowBound	2011	12353

Increase of standard deviation of coefficient b20 in the regression equation from 0.07dB to 1.4 dB significantly reduced reliability of biomass estimates, which is demonstrated by almost 10-fold increase of variance coefficient (CV) and very large confidential interval .

Frequency distribution functions for two versions biomass estimates





TS= 20 log L – 83, 61   sd=0.068 dB (1)

TS= 20 log L -83,41   sd=1.43 dB (II)

PROPOSALS FOR ACOUSTIC SURVEYS

IMPROVEMENT IN THE BALTIC SEA

The survey design

- The design of the Baltic acoustic surveys may be classified as systematic sampling.
- Participating vessels use different transect scheme designs : parallel transects , different types of zigzag and other.
- The processed data set of each vessel includes all acoustic samples on the transect and even data obtained in cross-points of zigzag transects are used.

However

- Processing of the acoustic data obtained by each participant of BIAS, is based on application of descriptive statistics algorithms (Descriptive analysis).

Such data processing is possible in the absence of any spatial correlation between the acoustic samples.

Analysis of correlation of acoustic samples should be made.

Only research of between-transects and transects correlation can indicate the validity of the above said approach to the acoustic data processing.

Otherwise, the methods of geostatistical analysis are more appropriate (Rivoirard et. al, 2000)

It seems that transect design should be taken into account in the method of BIAS acoustic data processing .

The skewed nature of acoustic data

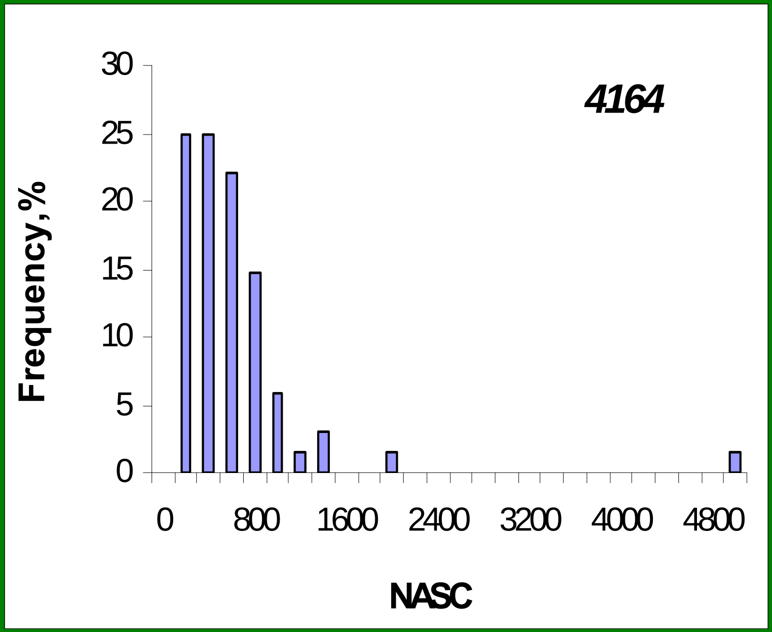
The acoustic data are processed by rectangles, based on standard statistic method of sampling analysis.

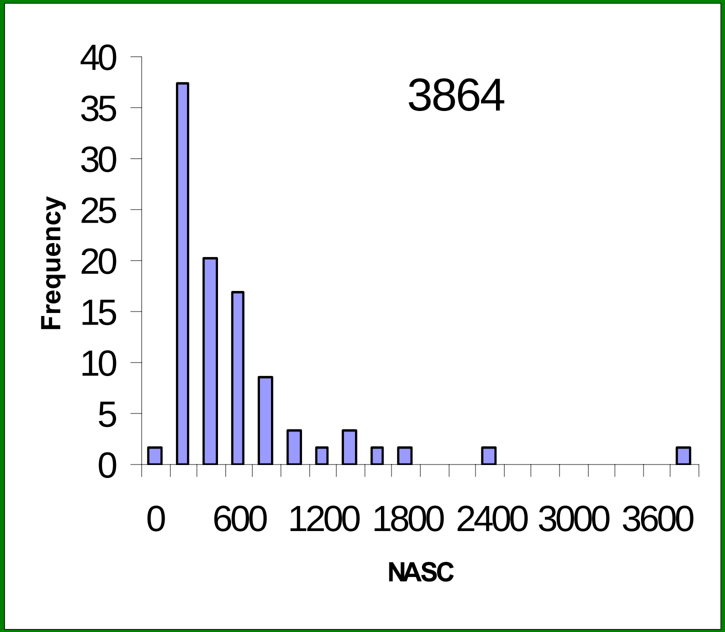
This approach is justified in the case of normal distributions of acoustic samples.

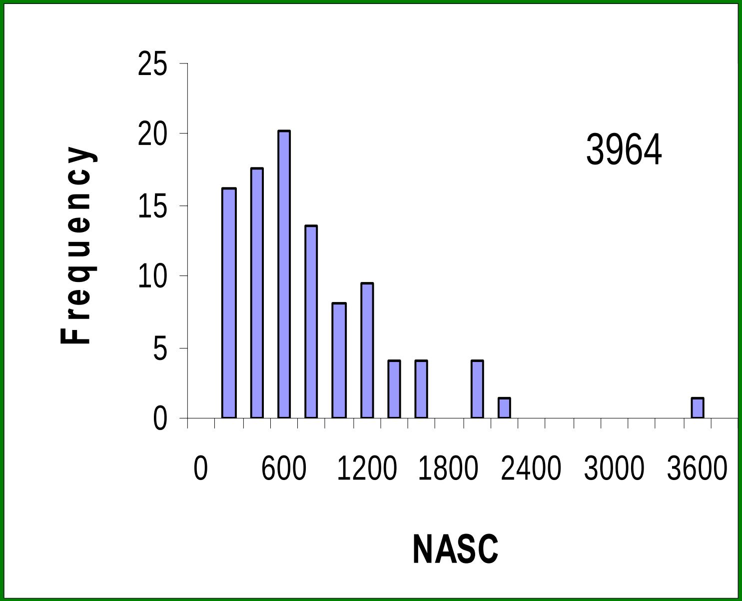
However,

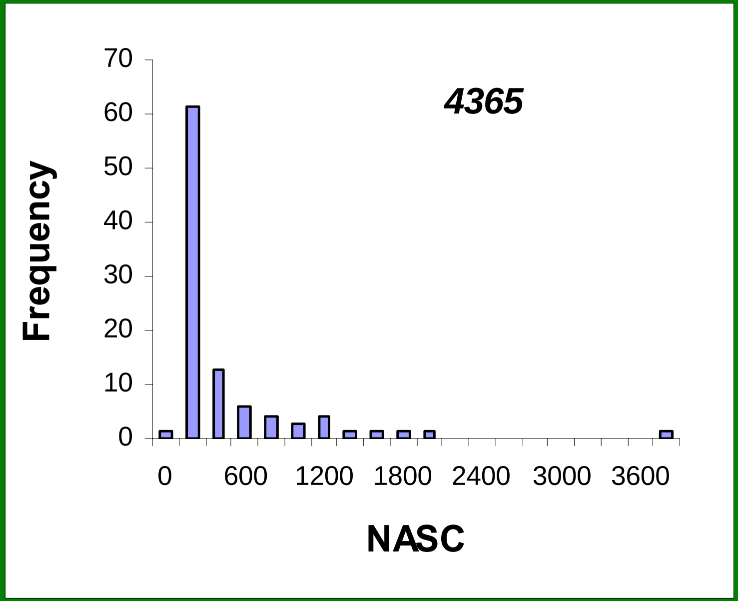
Typically, survey data have a highly positive skewed with few extreme values, and large proportion of zero and small observations.

As a result, estimates of mean density have large variances associated with the data processing when standard sampling formulas are used.


4164
Frequency,%
30
25
20
15
10
5
0
0
800
1600
2400
3200
4000
4800
NASC


3864
Frequency
40
35
30
25
20
15
10
5
0
0
600
1200
1800
2400
3000
3600
NASC


3964
Frequency
25
20
15
10
5
0
0
600
1200
1800
2400
3000
3600
NASC


4365
Frequency
70
60
50
40
30
20
10
0
0
600
1200
1800
2400
3000
3600
NASC

From of Subdivision 26 and 28 : 14 rectangles were tested

Log normal distribution of acoustic data was revealed for 5 rectangles

Sq	Arithmetical. mean				Delta-distribution			
	Mean	St.deviation	Confid. interval (90%)		Mean	St.deviation	Confid. interval (90%)	
3964	712,3	66,7	587,1	853,4	757,4	76,6	616,1	915,8
3965	94,5	51,2	26,2	208,8	87,7	50,2	30,2	220,6
4165	97,2	24,0	55,1	147,8	132,4	50,9	59,5	252,1
4465	254,5	26,1	209,2	307,6	264,1	29,3	213,3	324,9
4364	175,2	40,4	109,3	262,9	178,0	40,5	111,6	271,6

It is proposed to research the skewed nature of acoustic survey data distribution derived from Baltic surveys and select an effective estimator applying advance processing methods, which account for the distribution pattern (ICES WGSAD, 2004; Folmer and Pennington, 2000), (Pennington, 1983).

The problem of combination of control hauls

First aspect

- **Availability only two stations in some rectangles prevents the combination of control hauls and estimation of variances in analysing biological data.**
- **The latter does not allow to assess uncertainty of biomass and abundance estimates in view of the spatial variability of species composition and length structure of fish species in the survey area.**

The problem can be solved by means of changing the biological data averaging method. It seems that application of the geostatistical approach (Petitgas et al, 2003) is preferable

Second aspect

To correct equations recommended by Manual for International Acoustic Surveys (Report WGBIFS, 2000, 2003)

The species frequency f_i of species i :

$$f_i = \frac{1}{M}\sum_{k=1}^{M}\frac{n_{ik}}{N_k}$$

Cochran,1988

Jenssen 1978

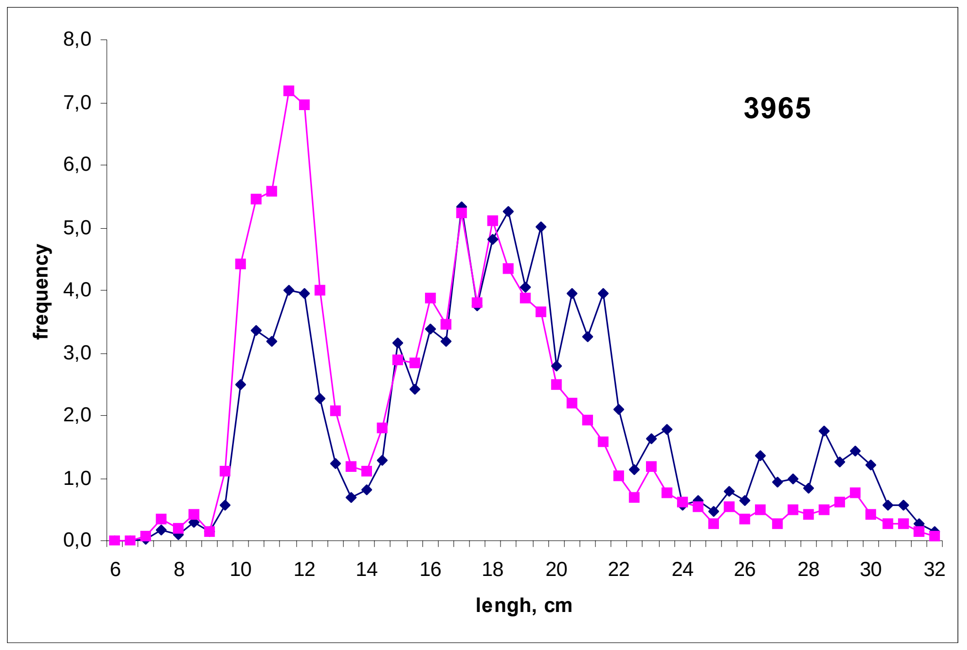
$$f_i = \frac{\sum_{k=1}^{M} n_{ik}}{\sum_{k=1}^{M} N_k}$$

The length frequency f_{ij} in the length class j as the mean over all M_i trawl catches containing the species

$$f_{ij} = \frac{1}{M_i}\sum_{k=1}^{M_i}\frac{n_{ijk}}{N_{ik}}$$

$$f_{ij} = \frac{\sum_{k=1}^{M_i} n_{ijk}}{\sum_{k=1}^{M_i} N_{ik}}$$

Examples

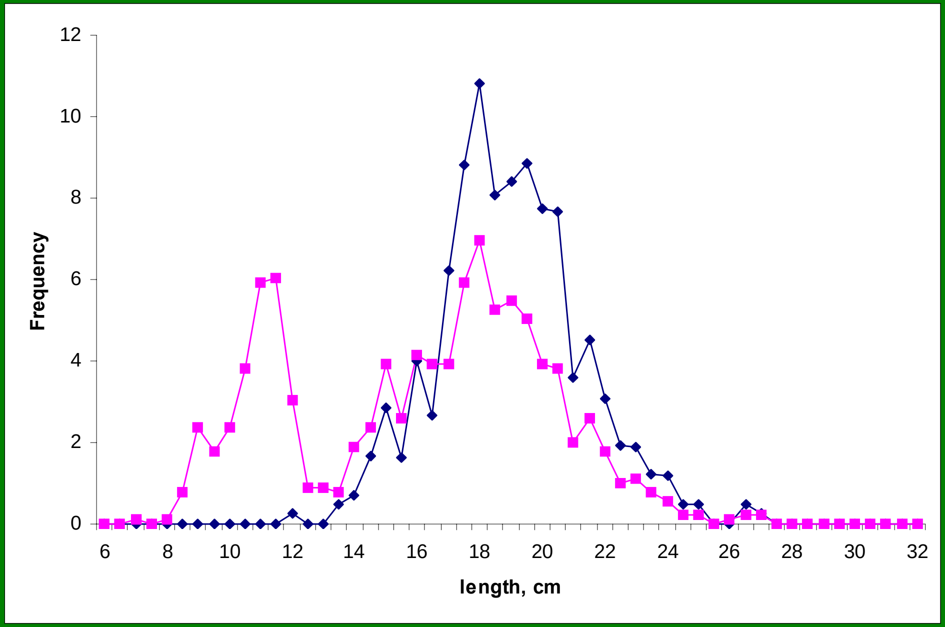


TS=- 46,68 dB

TS= -45,57 dB

TS –TS =1,11 dB

Ratio of density 1,29

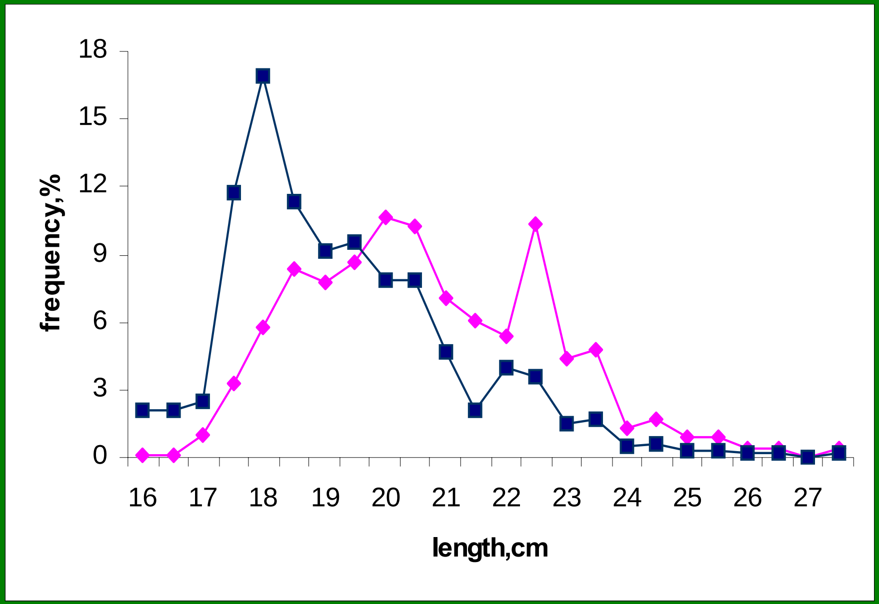


TS=- 47,06 dB

TS= -45,64 dB

TS –TS =1,44 dB

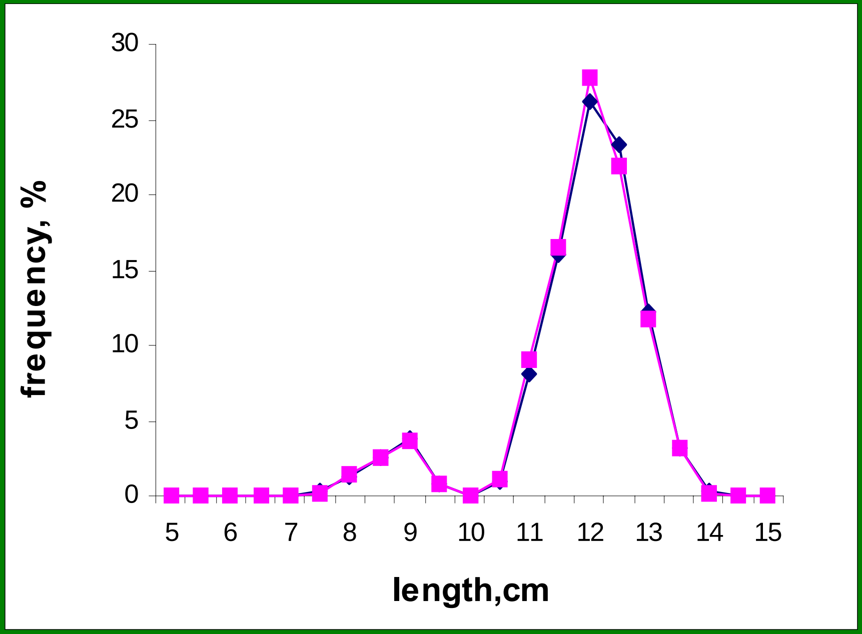
Ratio of density 1,33



TS=- 44,8 dB

TS=- 45,3 dB

Ratio of densities 1,13



Ratio of densities 1,005

TS in situ measurements

To speed up the estimation of the target strength equation parameters for the Baltic fishes.

Using ЕК500 and SonarData Echoview software, to conduct TS measurements *in situ* and adjust equations TS=F(L), used currently in surveys.

The adjusted TS regression should be accompanied by the statistical characteristics of its parameters

Uncertainty assessment for acoustic biomass estimates

- **To introduce the procedure of biomass and abundance uncertainty assessment into the Baltic survey methods.**
- **To analyze the basic sources of uncertainty of the BIFS survey results, taking into account algorithms of survey data processing and recommendations of ICES WG.**

Combination survey indices from different vessels

To modify the procedure of pooling the acoustic estimates by different vessels , taking into consideration their variances in the areas, where the surveys fulfilled were overlapped.