Evaluating the uncertainty of abundance estimates from acoustic surveys using geostatistical conditional simulations

Mathieu Woillez[a,*], Jacques Rivoirard[a], Paul Fernandes[b]

[a] *Centre de Géostatistique de l'Ecole des Mines de Paris, 35 rue St Honoré, 77300 Fontainebleau, France*
[b] *FRS Marine Laboratory, PO Box 101, 375 Victoria Road, Aberdeen, AB11 9DB, UK*

Abstract

Linear geostatistics, i.e. geostatistics based on the variogram, provide methods for estimating a spatially distributed abundance with its estimation variance. However, in complex situations, such as combining the data collected during acoustic surveys (i.e. acoustic backscattering, fish length and fish age), these methods are limited in their ability to combine the different sources of variability. As an alternative, geostatistical conditional simulations, which can reproduce the spatial variability of a variable, are particularly helpful in that they solve the latter problem.
In this paper, we investigate the uncertainty of Scottish herring acoustic survey estimates using a specific multivariate model. This includes highly skewed distributions for the acoustic backscatter data and incorporates relations between depth, mean fish length and proportion at age. Conditional simulations, i.e. simulations which honour the data values known at the data points, are used to generate multiple realisations for acoustic backscatter, mean fish length and proportion at age. These are combined to produce multiple realisations of herring density over the sampled domain. All realisations are then used to provide the error structure for the global abundance estimate at age. The method is used to assess the uncertainty from acoustic surveys on herring in the example year of 2003, and more generally for each year during the periods 1989 - 1994 and 2001 - 2003 to track significant variations of abundance over the time series.

Key words: geostatistics, simulations, acoustic surveys, herring, uncertainty.

1. Introduction

Scientific surveys, principally either trawl surveys or acoustic surveys, are often performed in order to estimate abundances of either demersal or pelagic fish populations respectively (Gunderson 1993). This estimation requires the combination of spatially located data. An important point is the uncertainty of this estimation, at least that part of the uncertainty due to the spatial coverage of the survey (e.g. areas between trawl stations or acoustic transects). As the variables of interest are often spatially autocorrelated, geostatistics have been recognized as providing suitable methods to estimate fish abundance from survey data (e.g. Petitgas, 1993; Rivoirard et al., 2000). This technique allows for:

- the capture of the spatial structure of fish densities (through the variogram, covariogram, and either with or without a trend or drift) and to fit an appropriate structural model: this is the geostatistical structural analysis;
- the estimation of global abundance with its estimation variance: this is the variance of the error of estimation, and can be computed from the structural model when the estimator is an arithmetic or weighted average, without preferential sampling;

* Corresponding author. Tel: + 33 (0) 1 64 69 47 76; fax: + 33 (0) 1 64 69 47 05. *E-mail address:* mathieu.woillez@ensmp.fr

- mapping by interpolation using kriging;
- simulations, conditional on the data (conditional simulations are simulations which honour the data values, by contrast with non-conditional simulations which reproduce the geostatistical model but do not honour data).

In many cases, the estimation variance of the abundance can be computed in a straightforward manner from a single variogram model and then be used to summarize the spatial uncertainty of the abundance. In more complex cases however, geostatistical simulations are needed to provide the uncertainty. Simulations allow the reproduction of the spatial variability of the variable, and also the simulation of errors and the computation of estimation variances. They are particularly helpful to compute estimation variances or probability when different sources of variability must be combined, e.g. acoustic backscatter together with biological data in acoustic surveys.

To date there have only been a few examples where overall survey variability has been evaluated. Measures of uncertainty have been based on sampling error (Rose et al., 2000) and survey timing, detectability, species composition, target strength, calibration coefficients and missing strata (O'Driscol, 2004). Rare are the cases where variances of abundance estimates are evaluated for acoustic surveys (e.g. Demer 2004). Gimona and Fernandes (2003) have attempted an approach to build geostatistical conditional simulations on Scottish herring (*Clupea harengus*), but the simple approach they used was not able to deal with the specific distribution of acoustic values, with many zeroes and a few high values, and resulted in bias. An alternative, more considered approach is proposed in this paper, again using the example of the Scottish component of the North Sea herring acoustic survey (Bailey et al., 1998).

2. Material and methods

2.1 Data

Acoustic surveys have been carried out in the northern North Sea (western half of ICES division IVa) in midsummer of each year since 1979 on the pre-spawning concentration of autumn spawning herring. The investigation here used data collected from nine years by the research vessel Scotia around Orkney and Shetland (1989-1994 and 2001-2003). This survey is part of a larger international survey for North Sea herring (Bailey et al., 1998). The result of the larger survey, is used in the assessment process, which ultimately aims to determine the total biomass, total numbers and numbers at age of the North Sea herring stock (Simmonds, 1996).

The Scottish survey is made up of transect lines that cover a domain defined according to ICES rectangles (Figure 1a). Parallel transect spacing is variable depending on the historically perceived levels of abundance of the area being sampled: 30, 15 or 7.5 nautical miles (nmi) are chosen. Acoustic backscatter data were recorded using a Simrad EK500 echosounder operating at 38 kHz. The echosounder system was calibrated on each survey according to recognised procedures (Foote et al. 1987). Trawl hauls were taken regularly to assist in the identification of the acoustic backscatter as detected as echo traces by the echosounder and to collect biological data such as fish length and fish age (Figure 1b and 1c). Echo traces were allocated to the appropriate fish species by visual scrutiny of the echogram (Reid et al., 1998). Outputs of the scrutinised acoustic backscatter data were values of the nautical area backscattering scattering coefficient (equivalent to $s_A/4\pi$, in $m^2.nmi^{-2}$) attributed to herring for an equivalent distance sample unit (EDSU) of 2.5 nmi.

2.2 Global geostatistical model

The simulation was based on a global geostatistical model combining acoustic backscatter (recorded along transects), mean length and proportions at age (measured at trawl stations), that were developed for the Scottish herring using the data from the surveys 1989-1994 (Rivoirard et al., 2000). In the present paper the model has been extended to include several more recent years. The essential characteristics of this model are now recalled. In this model, the mean

length increased with bottom depth and was mapped using kriging with depth as external drift. Moreover, as the spatial distribution of fish length is rather stable over years, the map of mean length for a given year was improved by using the length data from the other available years within an adequate three-dimensional model, where interannual variability was incorporated as a time component of the model (a nugget). Although the spatial distribution of age varied from year to year, it was not independent from that of length. For a given year, the cumulative proportion at age was closely related to the mean length through a logistic relation, and was mapped by kriging using this logistic regression on mean length as external drift and a residual structure with no spatial discontinuity. Maps of proportions at age were then obtained by differences between the maps of cumulative proportions. The acoustic backscatter was mapped by kriging. A map of the number of individuals was obtained by combining the acoustic map and the length map, using the target strength to length relationship for North Sea herring (ICES 2004). Finally, this is disaggregated into maps of numbers at age using the maps of proportions at age.

In summary, maps of acoustic backscatter, mean fish length (dependent on depth), and proportions at ages (dependent on mean fish length) were combined in order to provide the total abundance and abundances at age. In the present paper, repeated conditional simulations were used instead of kriging: this gave repeated simulations of abundances, conditional on the different types of data, and estimates of uncertainty of the estimated abundances. Kriged estimates of the different variables (acoustic backscatter, fish length and proportions at age) were also obtained to verify that the average of the simulations was close to the kriged estimates. We will now consider the method used for the conditional simulations, and afterwards the particular case of the acoustic backscatter.

2.3 Conditional simulation

A conditional geostatistical simulation of a Gaussian random function model can be classically done as follows (Chiles and Delfiner, 1999). First, a non-conditional simulation of the random function model Z(x) is performed. This is simply a realization of this random function, say ZNCS(x), which reproduces the variability of the model. The turning bands method has been used to generate the non conditional simulation. Then the simulation must be conditioned, in order to honour the data values. Conditioning a simulation, i.e. making a conditional simulation from a non conditional simulation, can be done relatively easily using kriging in the Gaussian case. Let Z*(x) denote the kriged estimator of Z(x) at any point x. One can write:

$$Z(x) = Z^*(x) + [Z(x) - Z^*(x)] \qquad \text{true value = kriged estimator + kriging error} \qquad (1)$$

The kriging error is unknown since Z(x) is not known. However, in the Gaussian case, this error is spatially independent from the kriged process. The idea is to mimic such an error by using a non conditional simulation, known everywhere:

- first, generate a non conditional simulation ZNCS(x),
- krige it from its values at the data locations only: ZNCS(x)*,
- obtain the simulated error = simulated value – kriged value = ZNCS(x) - ZNCS(x)*,
- add this error to the kriging of the real variable:

$$\begin{aligned} Z_{CS}(x) &= Z^*(x) + [Z_{NCS}(x) - Z_{NCS}{}^*(x)] \\ &= Z_{NCS}(x) + [Z(x) - Z_{NCS}(x)]^* \end{aligned} \qquad (2)$$

As kriging is an exact interpolator (reproducing the data value at a data point), this is indeed a conditional simulation: by construction it reproduces the variability, and goes through the data points. Thus conditioning a Gaussian simulation is straightforward once the values of the Gaussian variable are known at the data points. The Gaussian assumption is fundamental. Very often the Gaussian assumption is not acceptable on the original variable (e.g. a skewed histogram), but is acceptable after a transformation, or anamorphosis: thus the conditional simulation is made on the Gaussian variable, and then backtransformed to have a simulation of

the original variable. However in the case of kriging with external drift, the Gaussian conditional simulation method still holds, with the Gaussian assumption reduced to the residual from the drift.

We have considered that the Gaussian conditional simulation technique could be applied with no transformation in the case of mean fish length and cumulative proportion variables, since their distribution is close to Gaussian and we work implicitly with the residuals from drifts. However, the method is not sufficient for the acoustic variable: because of the skewness of the distribution, a transformation is necessary, but the values cannot be directly transformed into Gaussian values because of the high proportion of zeroes.

2.4 Simulation of the acoustic variable

We suppose that the acoustic variable, say Z(x), conforms to a Gaussian anamorphosed model, i.e. it stems from a standard Gaussian random function Y(x) using a Gaussian anamorphosis Φ: $Z(x) = \Phi(Y(x))$. Knowing the anamorphosis is equivalent to knowing the histogram of Z, and we have $F(z) = P(Z(x) < z) = P(Y(x) < y) = G(y)$. The anamorphosis allows for the association of any value Y the value Z with the same cumulative probability. The reverse holds if the anamorphosis can be considered to be inverted. This is generally the case when Z takes positive values with no spikes. However, this is not the case for the zeroes, when the histogram of values of Z presents a spike of zeroes. For instance, if the values of Z have 50% zeroes, any negative value for the standard Gaussian Y can be associated to a zero value of Z. More generally, if the values of Z have a proportion of zeroes equal to p_0, any value of Y lower than yc can be associated with a zero value of Z, where the threshold yc is determined by the proportion of zeroes: $p_0 = P(Z = 0) = P(Y < yc) = G(yc)$

Hence it is not possible to transform directly the original zero data values of Z into Gaussian values. Since the Gaussian values of Y at data points are not all determined, the above conditional simulation method cannot be directly applied, and moreover the variogram of the Gaussian variable to be simulated cannot be directly computed. So some adaptations are needed.

First the values of Y at data points where Z is zero should be informed. Within the above Gaussian model, the Y values at all data points should conform to the variogram of Y (we will see later how to determine this variogram). A Gibbs sampler (Geman and Geman, 1984) was used to simulate iteratively the Y values at data points where Z is zero, conditional on their being lower than y_c and on the Y values at the other data points (Freulon, 1992; Freulon 1994; Lantuejoul, 2002). Once Y are informed at all data points, the classical conditional simulation method is applicable. Finally, a backtransformation of Φ is applied to obtain the conditional simulation of the raw variable.

Secondly, as seen above, the variogram of the Gaussian variable Y to be simulated cannot be computed directly because the values of Y cannot be determined directly at data points where Z is zero. However the variogram of any transformation of the Gaussian variable is a function of the variogram of Y. This can be used to determine indirectly the variogram of Y from the variogram of a transformed variable known at all data points. For this we have considered the "lower-cut" Gaussian variable:

$$Y^+ = yc.1_{Y<yc} + Y.1_{Y>yc} \qquad (3)$$

This transformed variable, equal to Y where Z is positive and to the continuous limit yc where Z is zero, is known at all data points.

3. Results

The different simulation steps are now presented in details for one specific year (2003). We will begin with the acoustic backscatter. This variable was characterised by a highly skewed distribution with a large number of small values, about 30% zero values, and only a few large values (Table 1). The Gaussian anamorphosis function was determined by plotting the raw data values against normally distributed values (Fig.2). This function was fitted using a spline

function. It was used at the end of the simulation process for the back-transformation of the Gaussian simulation into a raw realisation.

The values of the Gaussian variable were not known at data points where the acoustic backscatter is zero, but the values of the lower-cut Gaussian variable were known at all data points. The experimental variogram of the lower-cut Gaussian is presented in Figure 3. It was indirectly fitted by considering that the Gaussian variable had a nested structure with a nugget of 0.225, a spherical component of 0.375 with a range of 12 nmi and another spherical component of 0.425 with a range of 50 nmi. The models fitted for the others years also had nested structures with a nugget component and, most of the time, two spherical components; a short range one and a long range one (Table 2).

The simulation process of acoustic backscatter was developed in two parts. First the Gibbs sampling was used to simulate values of the Gaussian variable at the data points where the acoustic backscatter was zero (Fig 4). Then, using the classical conditional simulation method, a realisation of the Gaussian variable was produced (Fig 5a), which was finally turned into a raw acoustic backscatter realisation (Fig 5b).

Fish mean length and fish proportion above age were simulated more easily, as these were variables that were close to being Gaussian. The simulation process was a classical conditional simulation, with respect to the geostatistical model developed for the two variables, i.e. a kriging with depth as external drift with a time component for mean length residuals, and kriging with a function of kriged mean length as external drift for proportion above age residuals. A realisation of mean length is presented in Figure 6 and realisations of proportion above age are presented in Figures 7a and 7b. Then, a realisation at age was obtained as the difference between the realisations of proportions above successive ages (Fig 7c).

Finally, 200 conditional simulations of acoustic backscatter, mean length and proportions at age were produced for each year of the studied period. One can check that the conditional simulations respect the statistics of the data, as well as the means of the kriging estimates (Table 3). Total abundance was obtained by combining the first two simulated variables, whereas the abundances at age were obtained by disaggregating the total abundance according to simulated proportions at age. The histograms of numbers at age and total number obtained from the simulations are presented for year 2003 in Figure 8. The error distribution is rather symmetric and the kriging estimate lies within the distribution. The main objective of the simulation is to determine the error distribution and the uncertainty associated with the variables (Table 4). The coefficients of variation of the abundances at age (standard deviation of the simulated abundances divided by their mean) ranged from 5.9% to 46.2%. High CVs were found for extreme ages (age 1, age 8) when numbers were low. Otherwise CVs were mostly around 15%. The CVs of the total number were lower than CVs of numbers at age, around 10%.

4. Discussion

Concerning the methodology, Gibbs sampler has to be used with care. Indeed, it has been observed that the p.d.f. of the simulated data tends to a conditional Gaussian distribution as the number of iterations becomes very large. However, the rate of convergence toward equilibrium is not well known (Galli and Gao, 2001; Lantuejoul, 2002). In our case, the number of iterations was fixed to 1000. The resulting p.d.f. of the simulated values was checked to be Gaussian and the variogram was checked to be fitted by the model used for the simulation.

The interest of such results is that one can follow the numbers per age classes in time or cohorts with their associated uncertainty respectively in figures 9 and 10. Thus, Scottish North Sea herring decrease in total number from 1989 to 1994, as the distributions for the different years slightly overlap and their means show a constant decrease. Years 2001 to 2003 present a significantly higher level of abundance than years 1989 to 1994. However the increase of abundance from 2001 to 2003 may not be significant, as the distributions of years 2001 and 2003 overlap despite different means. In addition, age 2 has a higher abundance than age 1 (Fig.

10). The consistently higher CV's associated with age 1 abundances verify that this survey is not precise for this age group. This is not surprising given that younger fish tend to occur closer inshore, where the survey does not sample so well and younger fish often school at the surface where the acoustic apparatus cannot detect them. Age 2 might be more representative of the recruitment strength than age 1. Thus, the studied period presents two different levels of recruitment; weak in majority for the first period and higher for the last years with two strong year classes (Fig. 9). A follow-up of cohort with the associated uncertainties is also interesting in order to clearly identify the strength of the different cohorts and detect change in time. However, a complete time series would be needed to follow these changes properly.

This methodology appears to be very interesting to evaluate uncertainty of abundance estimates of acoustic survey. It could be extended on the whole North Sea Herring assessment process, probably by evaluating separately the uncertainty of abundance estimates of each surveys led in conjunction for that stock.

Acknowledgements

These developments have been done with the financial support of the European Union (program Fisboat - Fisheries Independent Survey Based Operational Assessment Tools, DG-Fish, 6thFP STREP, Contract No 502572). The analysis was carried out with an R library for geostatistical analysis (RGeoS) compiled by Nicolas Bez, Didier Renard, Jacques Rivoirard and Mathieu Woillez (Centre de Géostatistique, Ecole des Mines de Paris, Fontainebleau, France).

References

Bailey, M.C., Maravelias, C.D. and Simmonds, E.J. (1998). Changes in the distribution of autumn spawning herring (*Cupea harengus* L.) derived from annual acoustic surveys during the period 1984-1996. ICES Journal Marine Science,. 55: 545-555.

Chiles, J.-P. and Delfiner, P., 1999. Geostatistics: Modeling Spatial Uncertainty. Wiley, NewYork. 695 pp.

Demer, D.A. (2004). An estimate of error for the CCAMLR 2000 survey estimate of krill biomass. Deep-Sea Research Part II, 51: 1237-1251.

Foote, K.G., Knudsen, H.P., Vestnes, G., MacLennan, D.N. and Simmonds, E.J. (1987). Calibration of acoustic instruments for fish density estimation: a practical guide. ICES Cooperative Research Report, 144: 57 pp.

Freulon, X., 1992. Conditionnement du modèle gaussien par des inégalitées ou des randomsiées. Geostatistics Doctoral Thesis, School of Mines of Paris.

Freulon, X., 1994. Conditional Simulation of a Gaussian Random Vector with Non linear and/or Noisy Observations. In M. Armstrong and P.A. Dowd (eds.) Geostatistical Simulations, Kluwer, Dordrecht, 57-71.

Galli A., Gao, H., 2001. Rate of convergence of the Gibbs sampler in the Gaussian case. Mathematical Geology, Vol. 33, No. 6, 2001, 653-677.

Geman, S. and Geman, D., 1984. Stochastic relaxation, Gibbs distribution and the Bayesian restoration of images. I.E.E.E. Trans. Pattern Analysis and Machine Intelligence, 6, pp. 721, 741.

Gimona A. and Fernandes P. G., 2003. A conditional simulation of acoustic survey data: advantages and potential pitfalls. Aquatic Living Resources, Volume 16, Issue 3, Acoustics in Fisheries and Aquatic Ecology. Part 2, July 2003, Pages 123-129.

Gunderson, D.R. (1993). Surveys of fisheries resources, Wiley, New York, pp.

ICES. 2004. Report of the Planning Group for Herring Surveys. ICES C.M. 2004/G: 5.

Lantuejoul, C., 2002. Geostatistical simulation: models and algorithms. Springer. Berlin. 256pp.

O'Driscoll, R. L., 2004. Estimating uncertainty associated with acoustic surveys of spawning hoki (Macruronus novaezelandiae) in Cook Strait, New Zealand. ICES Journal of Marine Science, 61: 84-97.

Petitgas, P. 1993. Geostatistics for fish stock assessments: a review and an acoustic application. ICES Journal of Marine Science, 50: 285-298.
Reid, D.G., Fernandes, P.G., Bethke, E., A., C., Goetze, E., Hakansson, N., Pedersen, J., Staehr, K.J., Simmonds, E.J., Toresen, R. and Tortensen, E. (1998). On Visual Scrutiny of echograms for acoustic stock estimation. ICES CM 1998/J:3. pp.
Rivoirard, J., Simmonds, J., Foote, K.G., Fernandes, P. and Bez, N., 2000. Geostatistics for estimating fish abundance. Blackwell Science, Oxford. 206 pp.
Rose, G., Gauthier, S., Lawson, G., 2000. Acoustic surveys in full monte: simulating uncertainty. Aquatic Living Resources 13, 367-372.
Simmonds, E. J., Toresen, R., Corten, A., Pederson, J., Reid, D. G., Fernandes, P. G., and Hammer, C., 1996. 1995 ICES Coordinated acoustic survey of ICES divisions IVa, IVb, VIa, and VIIb. ICES C.M. 1996/H: 8.

Table 1. Descriptive statistics of the acoustic backscattering coefficient ($m^2.nmi^{-2}$) and mean length (cm) data by surveyed year (minimum, first quartile, mean, standard deviation (SD), third quartile, maximum, size and percentage of zeros).

Variable	Year	Min.	1st Quartile	Mean	SD	3rd Quartile	Max.	Size	% 0
Acoustic	1989	0.00	0.00	15.74	44.29	13.07	673.80	1096	41.14
	1990	0.00	0.00	17.21	36.28	16.59	384.20	1044	32.08
	1991	0.00	0.00	10.78	52.29	4.00	907.10	974	56.87
	1992	0.00	0.00	7.92	24.33	5.28	340.70	860	61.16
	1993	0.00	0.00	8.27	32.85	4.36	533.40	938	50.00
	1994	0.00	0.00	8.42	32.39	3.25	588.10	982	66.19
	2001	0.00	0.00	33.37	148.61	16.61	3531.00	957	54.75
	2002	0.00	0.00	30.03	94.92	21.26	1206.00	1002	47.10
	2003	0.00	0.00	24.81	68.50	23.49	1416.00	1109	32.37
Mean length	1989	9.381	21.16	22.86	4.59	26.06	28.76	21	
	1990	18.48	26.11	26.66	2.93	28.51	29.88	28	
	1991	24.65	28.47	28.94	1.43	29.88	30.93	24	
	1992	19.56	25.81	27.49	4.14	30.54	31.55	11	
	1993	19.81	27.62	28.86	3.10	31.34	32.58	32	
	1994	19.96	27.24	27.88	3.68	30.06	32.72	19	
	2001	23.23	25.18	26.28	1.63	27.25	30.13	42	
	2002	18.45	25.83	26.76	2.10	28.24	30.54	45	
	2003	22.25	24.70	26.19	2.11	27.52	30.08	39	

Table 2. Variogram models of the Gaussian variable of acoustic backscatter in 1989-1994 and 2001-2003, fitted indirectly on lower-cut Gaussian variable.

Year	Model components	Nugget effect	First component		Second component	
			Sill	Range (nmi)	Sill	Range (nmi)
1989	nugget + spherical + spherical	0.200	0.450	12	0.350	100
1990	nugget + exponentiel + spherical	0.150	0.750	5	0.100	50
1991	nugget + spherical + spherical	0.400	0.450	15	0.150	100
1992	nugget + spherical + spherical	0.350	0.300	6	0.400	40
1993	nugget + spherical + spherical	0.320	0.530	15	0.250	50
1994	nugget + spherical + spherical	0.300	0.400	10	0.300	35
2001	nugget + spherical + spherical	0.100	0.400	11	0.800	65
2002	nugget + spherical + spherical	0.200	0.250	10	0.725	55
2003	nugget + spherical + spherical	0.225	0.375	12	0.425	50

Table 3. Comparative statistics of acoustic backscattering coefficient ($m^2.nmi^{-2}$) and mean length (cm) for each surveyed year. Estimates considered here are the mean and variance of raw data weighted by surfaces of influence, the mean and variance of kriged values and the simulated values, where mean and variance have been computed for each simulation before averaging the whole over the 200 simulations.

Variable	Year	Weighted data		Kriged values		Average of the simulated values	
		Mean	Variance	Mean	Variance	Mean	Variance
Acoustic	1989	15.68	2145.11	15.68	807.16	15.66	2193.94
	1990	15.06	1071.93	15.27	268.65	15.39	1095.35
	1991	9.96	2355.44	10.58	226.18	10.99	2968.25
	1992	8.08	623.42	7.67	66.40	8.29	701.02
	1993	6.72	849.84	7.27	102.12	7.47	1018.77
	1994	7.25	772.76	7.25	94.09	7.79	903.08
	2001	24.85	11216.89	26.66	2683.37	24.99	7953.22
	2002	26.48	7615.14	27.38	1473.71	27.14	6665.76
	2003	23.87	4628.00	24.37	1232.26	24.89	4845.56
Mean length	1989	21.64	34.65	21.93	20.07	22.29	21.17
	1990	25.81	10.87	25.84	11.85	26.08	12.51
	1991	28.71	2.35	28.07	7.61	28.23	8.28
	1992	28.40	9.86	26.70	5.00	26.79	6.50
	1993	29.06	12.28	28.93	5.76	29.04	6.48
	1994	28.03	11.58	27.40	7.94	27.60	8.52
	2001	26.74	3.09	26.73	4.51	26.82	5.13
	2002	26.22	9.42	26.18	7.69	26.37	8.31
	2003	25.98	5.38	25.60	7.47	25.74	8.21

Table 4. Statistics of the simulations of numbers at age and total number (in millions) for each survey year.

Age	1989			1990			1991			1992			1993		
	Mean	Sigma	CV	Mean	Sigma	CV	Mean	Sigma	CV	Mean	Sigma	CV	Mean	Sigma	CV
A1	2860	312	10.9	1282	208	16.2	208	92	44.2	523	139	26.6	315	140	44.4
A2	415	40	9.6	262	25	9.5	76	21	27.6	802	99	12.3	234	63	26.9
M2	2262	206	9.1	1170	92	7.9	471	98	20.8	1149	136	11.8	750	119	15.9
A3	103	11	10.7	95	7	7.4	0	0	0	0	0	0	219	36	16.4
M3	1907	218	11.4	1421	94	6.6	661	107	16.2	534	61	11.4	604	81	13.4
A4	533	86	16.1	1200	82	6.8	1343	186	13.8	476	53	11.1	372	45	12.1
A5	159	36	22.6	538	42	7.8	1223	159	13	524	64	12.2	341	41	12
A6	75	20	26.7	207	19	9.2	516	74	14.3	262	34	13	487	59	12.1
A7	25	7	28	159	17	10.7	279	52	18.6	56	13	23.2	282	37	13.1
A8	26	12	46.2	54	8	14.8	137	29	21.2	19	6	31.6	78	12	15.4
A9P	11	4	36.4	35	9	25.7	118	26	22	22	8	36.4	56	10	17.9
Total	8375	700	8.4	6424	382	5.9	5032	612	12.2	4368	438	10	3739	460	12.3

Age	1994			2001			2002			2003		
	Mean	Sigma	CV	Mean	Sigma	CV	Mean	Sigma	CV	Mean	Sigma	CV
A1	512	126	24.6	241	99	41.1	453	93	20.5	833	233	28
A2	161	35	21.7	1421	333	23.4	290	56	19.3	1599	200	12.5
M2	2042	267	13.1	4430	516	11.6	1880	209	11.1	3037	286	9.4
A3	5	1	20	76	11	14.5	22	4	18.2	285	29	10.2
M3	532	63	11.8	1519	131	8.6	5406	492	9.1	2313	201	8.7
A4	144	20	13.9	1160	100	8.6	1379	123	8.9	3804	358	9.4
A5	93	15	16.1	1623	147	9.1	1042	103	9.9	563	63	11.2
A6	103	16	15.5	487	49	10.1	1577	172	10.9	522	67	12.8
A7	148	22	14.9	219	25	11.4	442	52	11.8	846	118	13.9
A8	98	14	14.3	178	23	12.9	233	31	13.3	221	34	15.4
A9P	61	10	16.4	132	21	15.9	197	29	14.7	292	46	15.8
Total	3898	443	11.4	11485	1087	9.5	12921	1125	8.7	14316	1176	8.2

Figure 1a. Proportional representation of acoustic backscattering ($m^2.nmi^{-2}$) of herring from the Scottish survey around Orkney and Shetland in July 2003.

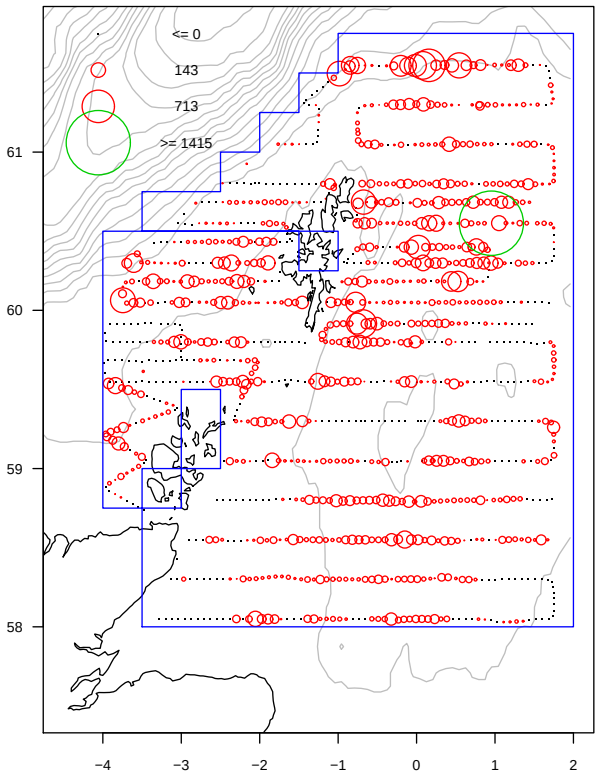

Figure 1b. Proportional representation of mean length (cm) of herring at trawl stations from the Scottish survey around Orkney and Shetland in July 2003.

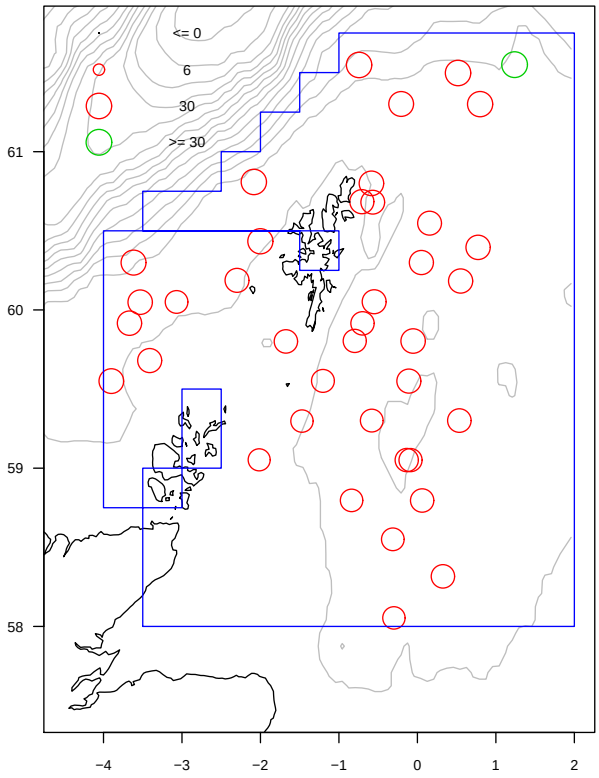

Figure 1c. Proportional representation of proportion at age 4 herring at trawl stations from the Scottish survey around Orkney and Shetland in July 2003.

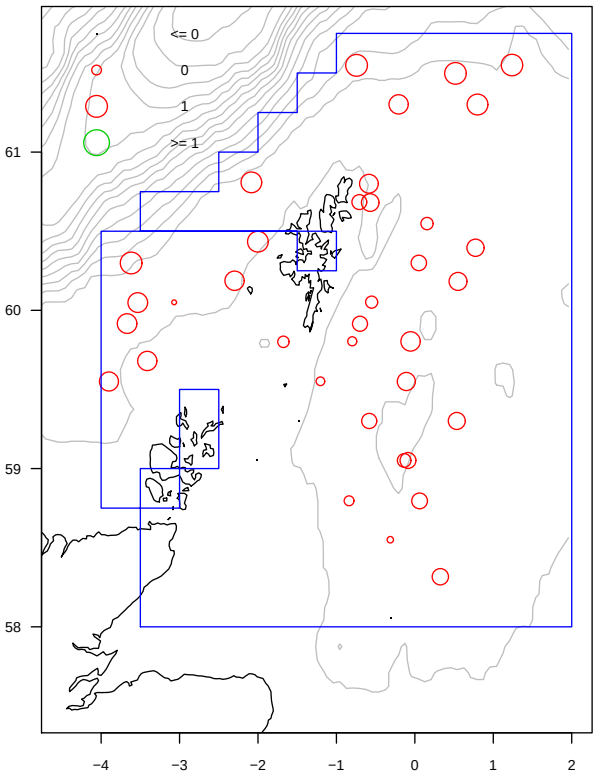

Figure 2. Gaussian anamorphosis, associating a raw value Z and a Gaussian value Y corresponding to the same cumulative probability. A spline function was used to model the Gaussian anamorphosis.

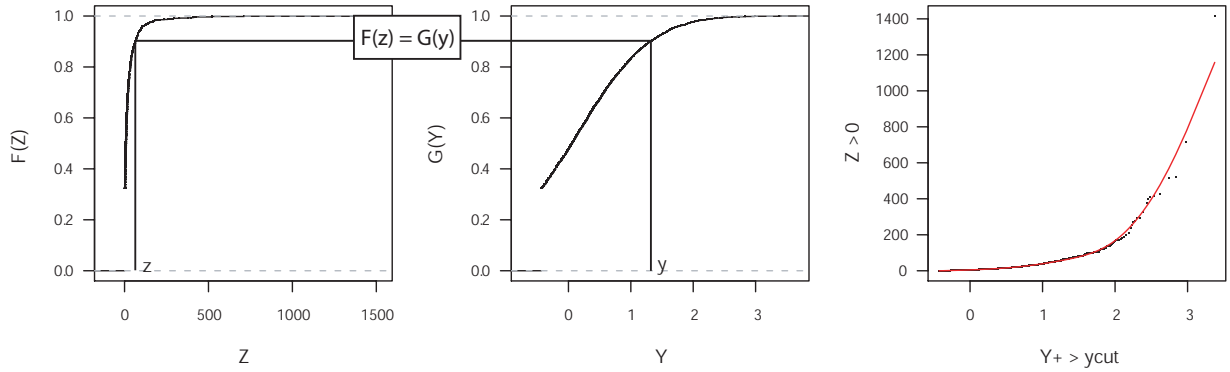

Figure 3. Experimental variograms of the lower cut Gaussian (Y+) with circles proportional to number of data and its indirect fitted model (dash line). The model of the Gaussian (Y) used for the indirect fitting is the solid line.

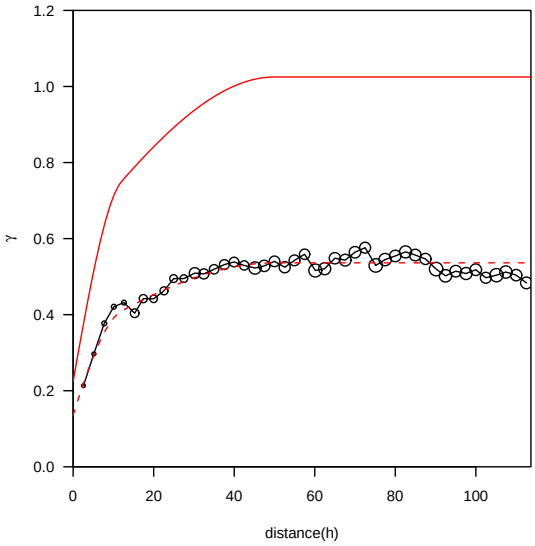

Figure 4. Histogram of Y+. Histogram of Y obtained by Gibbs sampling. Variogram of Y corresponding to the desired model used to build the left part of the histogram of Y.

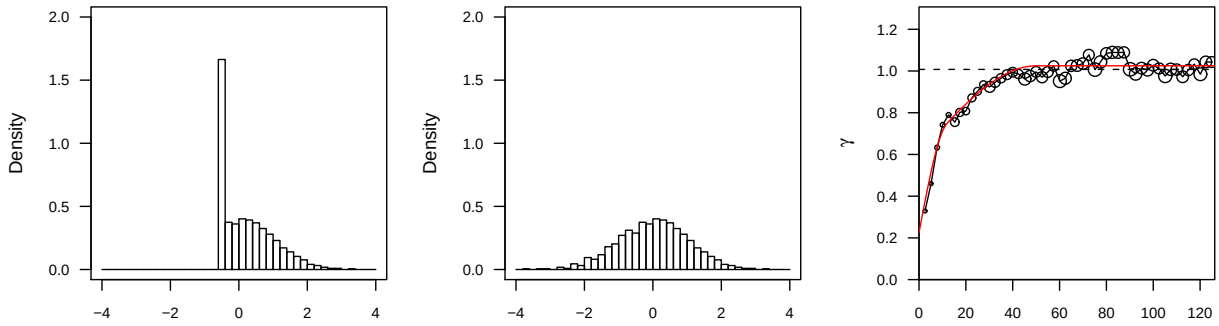

Figure 5a. A realisation, or conditional simulation, of the acoustic variable in Gaussian space. Only values above the Gaussian threshold yc corresponding to zero in the raw values have been displayed.

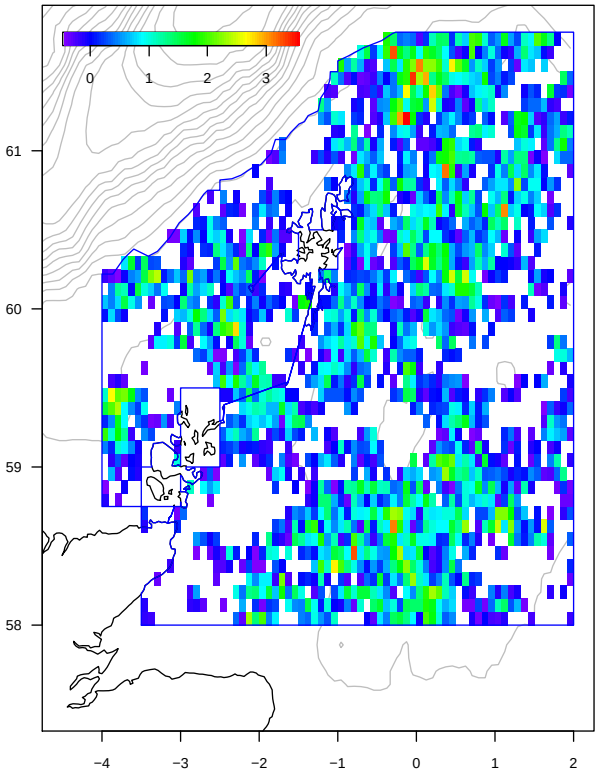

Figure 5b. A realisation, or conditional simulation, of the acoustic variable obtained after the reversed transformation. Only simulated values above zero have been displayed.

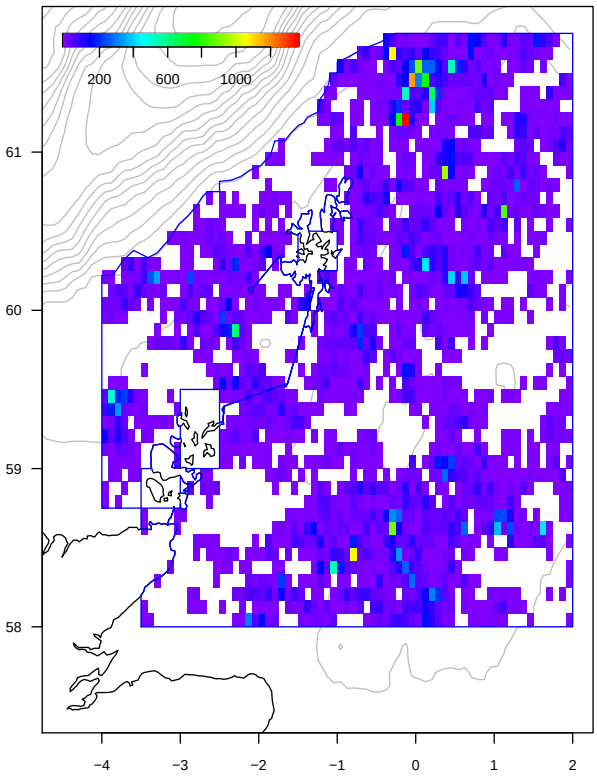

Figure 6. A realisation of mean length (cm) variable based on trawl data.

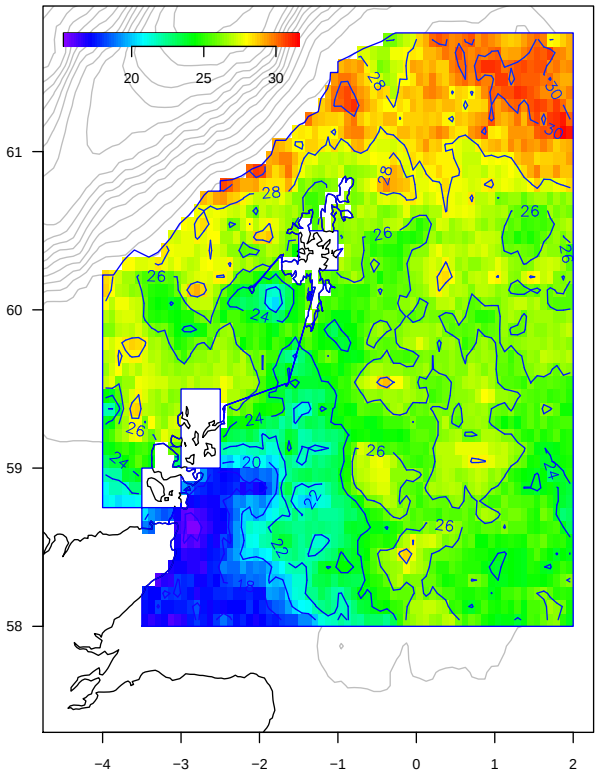

Figure7a. A realisation of proportion above age 3

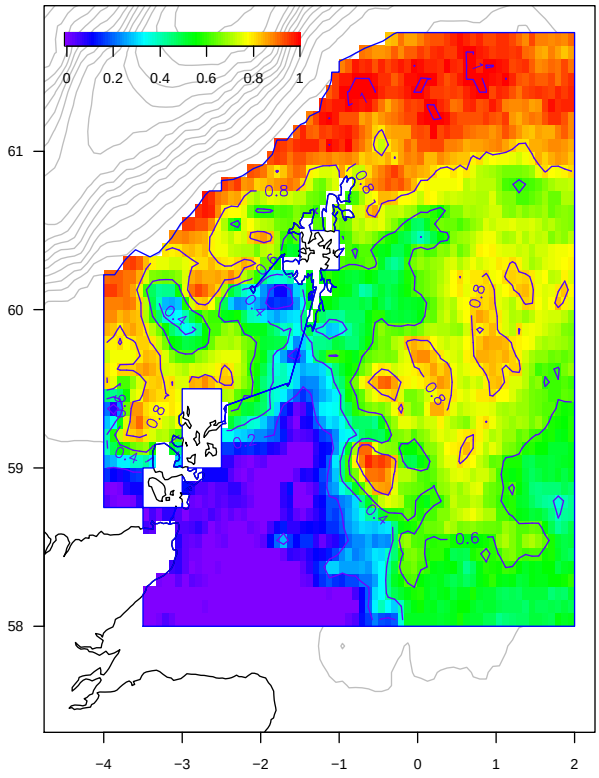

Figure7b. A realisation of proportion above age 4

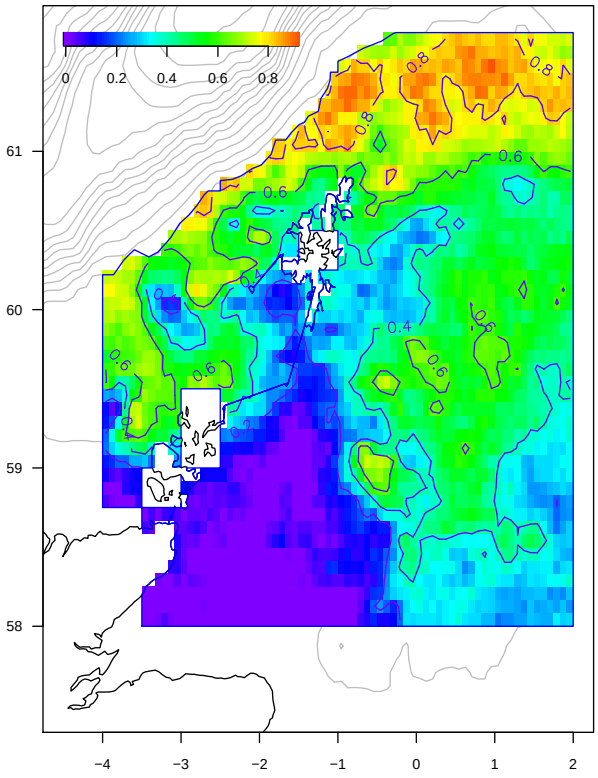

Figure7c. A realisation of proportion at age 3, obtained by difference between the two previous realisations.

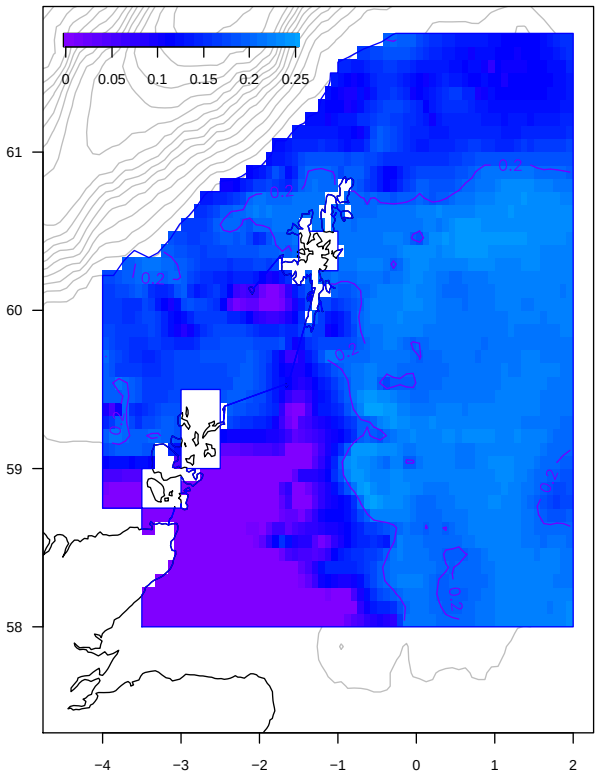

Figure 8. Error distribution of numbers at age and total number (millions), year 2003. Vertical lines are the kriged estimates for each distribution.

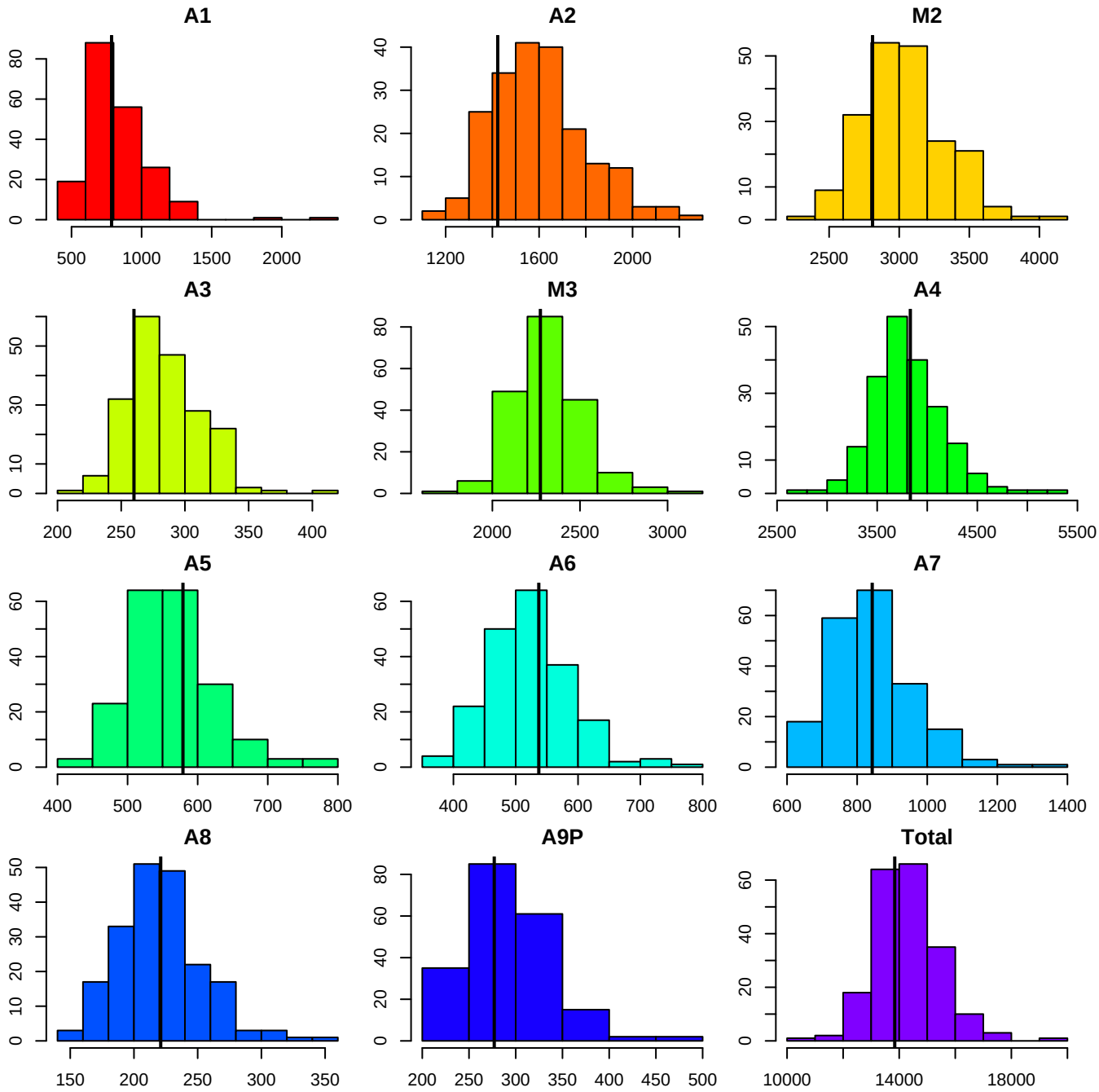

Figure 9. Boxplots of the different age classes and the total number (millions) along the time series.

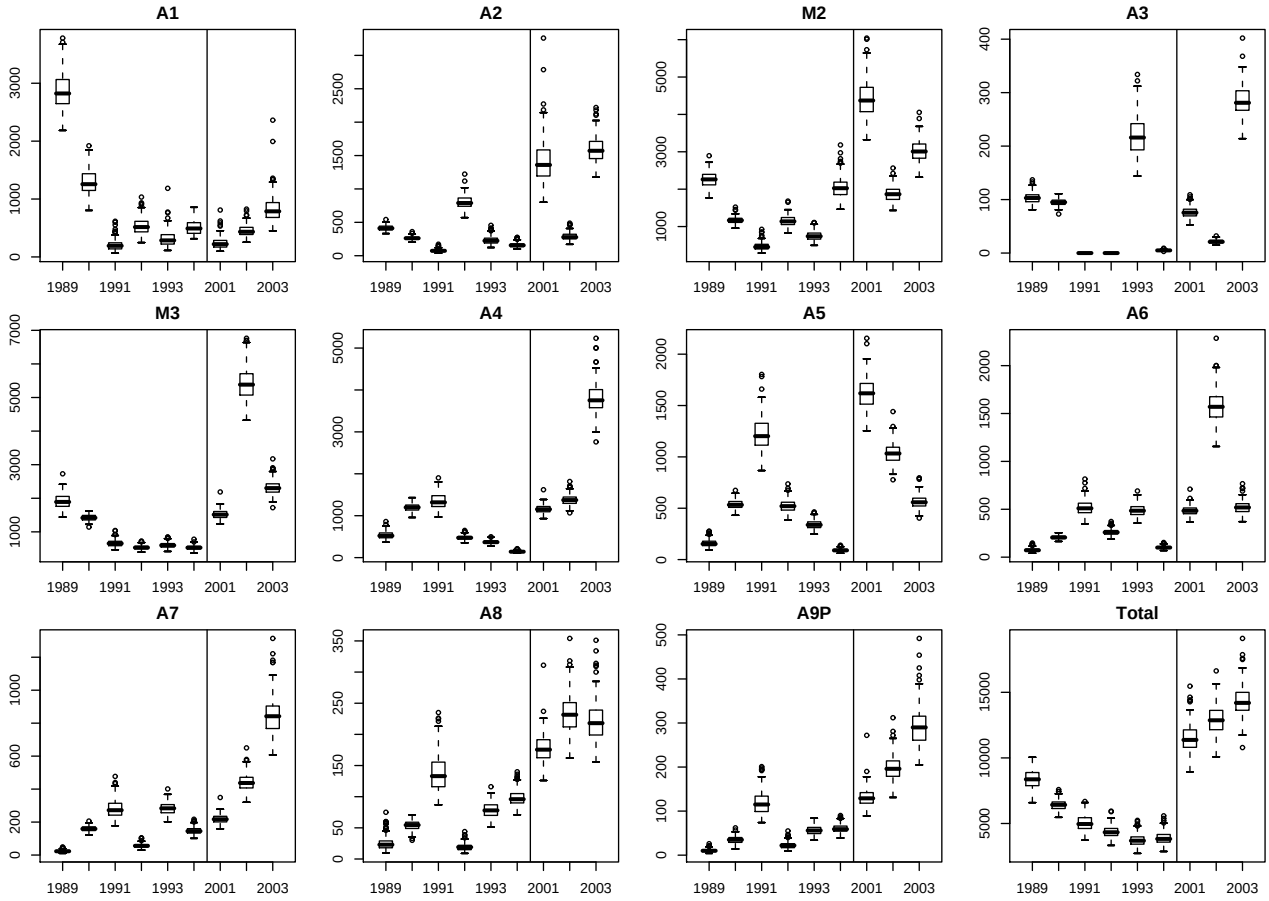

Figure 10. Boxplot of the available cohorts.