Bayesian methods for the estimation of spawning stock biomass from pelagic egg surveys

Iago Mosqueira Sánchez

A thesis submitted in fulfilment of requirements for the degree of Doctor in Philosophy and the Diploma of Imperial College

Renewable Resources Assessment Group
Imperial College
University of London

2005

Abstract

The diminishing captures experienced in recent years in many important fisheries have motivated a search for new exploitable species, and the need for more informative assessment methods. Small pelagics, such as Atlantic mackerel *Scomber scombrus*, have seen their importance increase over time, as the traditionally highest-valued stocks have become less abundant. Atlantic mackerel, for example, has turned in the last three years into the most valuable stock in British and Irish waters. The direct importance of these stocks is increased by their role in the oceanic ecosystem, and the possible effects of fishing on lower levels of the food chain (Pauly *et al.*, 1998).

Atlantic mackerel resources in the ICES area are assessed directly every three years by means of ichthyoplankton surveys and the application of the Annual Egg Production Method (AEPM) (ICES, 1999a). Egg production is surveyed over the whole of the spawning area and season, and then combined with fecundity data to arrive at a Spawning Stock Biomass (SSB) estimate (Gunderson, 1993). The method applied at present to these is extremely simple in its treatment of both the spatial and temporal dimensions, and does not take into account the uncertainty and bias present in the procedures involved (ICES, 2000).

Various methods are presented that allow for both improved, by estimating the possible bias incurred, and more realistic estimates, by quantifying the uncertainty present in the "traditional" estimator. Analysis of various sources of data have allowed the construction of probability distributions for the parameters involved in the estimation procedure. Bayesian methods (Gelman *et al.*, 1995) were applied to assess their impact in the final SSB estimate.

Spatio-temporal geostatistical estimates of egg abundance have been developed that show the potential for better use of the available information, and for future survey designs that would allow for more precise estimates. A Bayesian implementation of those estimators is presented that show the importance of different sources of uncertainty on the final estimate, especially in the variogram modelling step. Various variogram models are compared and a choice is made of the ost appropiate for the data under scrutiny. Block-kriging in a three-dimensional setting is applied using the result of the variogram modelling exercise. From this, estimates of SSB are generated and compared with those

previously obtained. The influence of these results on the assessment of the NEA mackerel stock is briefly explored.

Acknowledgements

First of all I would like to thank most wholeheartedly the invaluable help and advice received from my colleagues at RRAG and Imperial College, especially Claire Imrie, Catherine Michielsens, Shelley Clarke, Richard Hillary and Panayiota Apostolaki. Very special thanks go to Edzer J. Pebesma, author of *geostat*, for the time and effort put into adapting the program to my needs. Fréderic Bois, author of *MCsim*, kindly answered all my questions regarding his software.

Much was learned from researchers in the GBMAF project and the ICES Working Group on Mackerel and Horse Mackerel Egg Surveys, most notably Douglas Beare and Dave Reid in Aberdeen, and Peter Witthames at Lowestoft. Steve Coombs provided me with invaluable data on vertical distribution of mackerel eggs. The Centro de Supercomputación de Galicia, CESGA, provided computing time on their machines. Muitas gracias to Carlos Fernández for his help.

Finally, I would like to thank my supervisors Dr. Murdoch McAllister and Dr. Anna Korre for giving me this unique opportunity of learning and discovery.

Contents

List of Figures 8

List of Tables 11

1 Introduction 12

1.1 Problem statement . . . 13
1.2 Purpose and description of the research . . . 14
1.3 Ichthyoplankton and fish stock biomass estimation . . . 15
1.4 The Bayesian paradigm: modelling uncertainty . . . 17
1.4.1 Bayesian methods in fisheries assessment . . . 20
1.5 Geostatistical methods . . . 20
1.5.1 Use of geostatistics in fisheries assessment . . . 21
1.5.2 Bayesian geostatistics: developments and applications . . . 23
1.6 The Northeast Atlantic mackerel fishery . . . 25
1.6.1 Stock structure . . . 26
1.6.2 The European mackerel egg surveys . . . 27
1.6.3 Exploratory analysis of the dataset . . . 30
1.6.4 Assessment and management needs . . . 32
1.7 Thesis outline . . . 34

2 Review of the Annual Egg Production Method 38

2.1 Data collection . . . 38
2.2 Data management . . . 40
2.3 The AEP model . . . 41
2.4 Model parameters . . . 43
2.4.1 Depth . . . 43
2.4.2 Stage I length . . . 44
2.4.3 Water temperature . . . 44
2.4.4 Area . . . 44
2.4.5 Volume . . . 44
2.4.6 Fecundity . . . 46
2.4.7 Egg counts . . . 47

2.5 Model assumptions and procedures 49
2.5.1 Instantaneous sampling 49
2.5.2 Net speed and depth coverage 49
2.5.3 Use of t_{20} as egg development temperature 56
2.6 Bias model 58
2.7 Conclusions 63

3 Bayesian variogram modelling **68**
3.1 Estimation of experimental variograms 69
3.1.1 Data manipulation 69
3.1.2 Sample semi-variogram estimation 71
3.1.3 Three-dimensional space-time variograms 72
3.2 Variogram models 72
3.2.1 Valid models 72
3.2.2 Nugget model 73
3.2.3 Exponential model 73
3.2.4 Spherical model 73
3.2.5 The Matérn model 74
3.2.6 Nested models 74
3.3 Bayesian hierarchical modelling for spatio-temporal variograms 74
3.3.1 Selection of prior probability distributions 75
3.3.2 Likelihood function 76
3.3.3 MCMC integration 76
3.3.4 Convergence diagnostics 77
3.3.5 Bayesian model checking and sensitivity analysis 78
3.3.6 Use of Bayes' posteriors for model choice 78
3.3.7 Implementation of the Bayesian and geostatistical methods applied 79
3.4 Results 79
3.4.1 Sample variograms 79
3.4.2 Assessing the exchangeability assumption 79
3.4.3 MCMC Convergence 81
3.4.4 Prior sensitivity analysis 81
3.4.5 Model fits 81
3.4.6 Parameter estimation and uncertainty 82
3.4.7 Relative nugget 83
3.4.8 Model choice 85
3.5 Discussion 97

4 Monte Carlo block kriging of egg abundance **100**
4.1 Three-dimensional block kriging 100
4.1.1 Block kriging 101

4.1.2 Spatio-temporal kriging 101
4.1.3 Implementation of the kriging procedure 103
4.1.4 Search neighbourhood 103
4.2 Monte Carlo kriging estimation 105
4.2.1 Implementation details 106
4.3 Kriging estimates of North East Atlantic mackerel Annual Egg Production 107
4.4 Impact of model choice 108
4.4.1 Generation of simulated datasets 109
4.4.2 Model comparison using simulated datasets 113
4.5 Inferences on spawning dynamics from kriging 115
4.5.1 Centre of gravity of spawning 115
4.5.2 Timing of spawning 120
4.6 Discussion 120

5 Comparison of estimators and implications for assessment and management 123
5.1 Bias-corrected estimates of annual egg production 124
5.2 Estimates of spawning stock biomass from kriging-derived annual egg production 125
5.3 Factors not included in the analyses 126
5.3.1 Survey design and estimator choice 130
5.4 Effect of uncertainty on stock assessment 131
5.4.1 Assessments conducted by ICES Working Group 132
5.4.2 Absolute or relative indices of SSB 132
5.5 Final discussion 134
5.5.1 Estimator performance 135
5.5.2 Management implications 136
5.5.3 Future research 136

A Notation 139

B Acronyms 140

References 141

List of Figures

1.1 Flow diagram of the process behind the AEPM 16
1.2 Catches of Northeast Atlantic mackerel between 1969 and 2002 26
1.3 Catch-at-age of Northeast Atlantic mackerel in 2002 27
1.4 Sampled locations per year 29
1.5 Standard survey areas used in ICES mackerel egg surveys 30
1.6 Three-dimensional representation of the spatio-temporal distribution of samples by year. 31
1.7 Egg counts and sample sizes per month and year 33
1.8 Time cumulative egg abundance by year. Abundances expressed as a percentage of the highest historical egg count. 34
1.9 Egg abundances as raw egg counts vs. Julian day, latitude and longitude, and by year. 35
1.10 Egg abundances per year 36

2.1 Catch curve method. 42
2.2 Distribution of normalised residuals around mean fecundity estimates for Atlantic mackerel 47
2.3 Results of the 1998 (a) mackerel and (b) horse mackerel egg exchange exercise. Distribution of the deviations from the mean proportion of stage I eggs in sample ($n = 10$). 48
2.4 Probability distribution of t_n/t for Atlantic mackerel. 50
2.5 Spatial location of egg profiles for the three central months of the survey . 51
2.6 Plots of the relationship between depth in meters where most eggs are concentrated, ($\bar{d}$), and the percentange of eggs found below 200m (N_{200}), against latitude (ϕ), bottom depth (B), Julian date (J), and egg numbers in the sample N, in log scale. 54
2.7 Dsitribution of values of ϵ_d. 55
2.8 Correlation plots between the explanatory variables in the t_{20} and depth coverage analyses. 58
2.9 Model of $\log(\epsilon_t + 1)$ change with date. 59
2.10 Estimation procedure for q, the constant of proportionality. 60

2.11 Distribution of the ratio of bias corrected (D_c) to traditional DEP (D) estimates. 61
2.12 Distribution of the ratio between bias corrected, D_c, and traditional, D, daily egg production estimates. 63
2.13 Distributions of the ratios between depth bias-corrected, and temperature bias-corrected, against the traditional daily egg production estimates . . . 64
2.14 Distribution of bias arising from the t_{20} procedure, ϵ_t, from 500 iterations on the 1998 ICES survey data, and on the four depth and time strata. . . 65
2.15 Distribution of bias from the 200m sampling limit from 500 iterations on the 1998 ICES survey data. 65

3.1 Example of data transformation with both methods 70
3.2 Variogram models applied in this study: exponential, spherical and nested combination of both (a), and Matérn (b). The Matérn model in panel b is shown for a number of values of the κ shape parameter and for the same combination of range and sill. 73
3.3 Sample variograms for mackerel egg abundances 80
3.4 Example MCMC chain traces of the nested model hyperparameters for the nested model. 82
3.5 Posterior probability distributions for ϕ_x and C obtained with divergent prior probability distributions. 83
3.6 Variograms obtained with the Matérn model 85
3.7 Variograms obtained with the exponential model 87
3.8 Variograms obtained with the spherical model 89
3.9 Variograms obtained with the nested model 91
3.10 Posterior distributions of the ϕ_x parameter 92
3.11 Posterior distributions of the sill C 93
3.12 Posterior distributions of the C_0 parameter 94
3.13 Posterior probability for the nugget C_0 of the Matérn model in 1977. . . . 94
3.14 Joint posterior distributions of ϕ_x and s. 95
3.15 Joint posterior distributions of ϕ_z and C. 96
3.16 Example relative nugget posteriors 96

4.1 Discretisation procedure in block kriging 102
4.2 Center points of the 15x15 nautical mile blocks defined for kriging. 104
4.3 Maps of weekly egg production estimated by block kriging using different temporal search radii 105
4.4 Estimates of AEP and credibility intervals for each model 109
4.5 Probability distributions of AEP kriging estimates 110
4.6 Weekly egg production for all models and for 1986. 111
4.7 Weekly egg production for all models and for 1998. 112

4.8 Sample variograms obtained from the simulated datasets for 1998 113
4.9 Posterior pdfs for exponential model and the SGS-generated data 114
4.10 Position by week of the centroid of the egg abundance matrix 116
4.11 Kiriging-estimated weekly mackerel abundance in 1998 117

5.1 Estimates of bias-corrected AEP with credibility intervals 126
5.2 Probability distributions of the bias-corrected AEP 127
5.3 Probability distributions of SSB obtained 128
5.4 Median spawning stock biomass estimate and 95% credibility intervals . . 129
5.5 SSB estimates from ICA using kriging-based index of abundance 132
5.6 Estimates of SSB obtained by ICES WGMHSA 133
5.7 SSB trajectories from ICA with absolute or relative index 134

List of Tables

1.1 Examples of applications of the AEPM 18
1.2 Summary information of surveys undertaken in the NE Atlantic for mackerel eggs between 1977 and 2001. 28

2.1 Parameters in the AEPM model 44
2.2 Depths covered on each section of the egg and temperature profiles 57
2.3 Percentages of the daily egg production and sampling sites found on both sides of the 200m isobath. 62
2.4 Egg counts obtained during repeated sampling of a single square 66

3.1 Prior probability distributions for the variogram parameters. 76
3.2 Mean and CVs for the different hyperparameters under the four models tested. 82
3.3 Values obtained for the various model fit metrics and for each model. . . . 87

4.1 Parameter values used for the searching procedure in both dimensions. . . 105
4.2 Kriging estimates of NE Atlantic mackerel annual egg production 107
4.3 Comparison of kriging estimates of AEP with the ICES estimator 108
4.4 AEP kriging estimates using the simulated datasets 114

5.1 Bias-corrected AEP estimates 124
5.2 Estimates of SSB obtained from ICA using the kriging-based index of SSB 131

Chapter 1

Introduction

Atlantic mackerel (*Scomber scombrus*) is a migratory pelagic species widely distributed in the shelf regions on both sides of the North Atlantic. Species of the genus *Scomber* migrate in dense schools during the spawning season, and their eggs remain until hatching in the water column. Their spawn is generally less intense and restricted in time and space than that of sardines and anchovies, but is still concentrated and abundant enough to be detected and quantifiable by means of plankton sampling. Mackerel are determinate batch spawners, meaning that the potential fecundity of an individual on a given year is determined prior to the onset of spawning. Fisheries for small pelagics are becoming increasingly important in European waters, both due to increasing commercial value and diminishing catches from other, more traditionally sought-after species. Mackerel has become the most valuable stock for the U.K. fishing sector, both in terms of landings and market value (DEFRA, 2003).

Stocks of pelagic species (a term used from now on to refer strictly to small pelagics) are commonly assessed from both catch and fishery-independent data, through acoustic, trawl and planktonic egg surveys (Gunderson, 1993). Mackerel (*Scomber scombrus*) is being assessed by means of egg production methods in European waters triennially since 1977, when the first pelagic egg surveys were carried out on the stock off the Western European coast, from Biscay to Shetland. The survey was extended in 1995 to the Southern stock (nowadays called Southern component) in Spanish and Portuguese waters. From 1989 horse mackerel (*Trachurus trachurus*) eggs have also been sampled, and similar methods are employed for its assessment (ICES, 2000).

The methods applied to the analysis of these data have not changed much over the years, although improvements have taken place in survey equipment and methods. Both stocks are assessed by means of the Annual Egg Production Method (AEPM), that relates egg production in the spawning season and female fecundity to stock biomass (Gunderson, 1993). The growing dataset for these stocks has probably not been fully exploited, although recent publications have attempted to fill this gap (Beare & Reid, 2002). Despite the accumulating knowledge provided by new surveys and other studies, each survey year

is analysed without formal consideration of previous years results. Although spawning dynamics are known to change with environmental and other factors, each survey might be considered as a realisation of a common process: the annual migration and spawning cycle of a fish population. There is also growing concern among researchers and managers about the consequences of estimation errors in some problematic parameters, and of the untested assumptions implicit in the Annual Egg Production model applied.

It is the intention of this study to review the existing methodology for the estimation of stock biomass from egg survey data, identifying any significant source of error and/or bias, and to elaborate on two methods for improving the final estimates: (1) the use of the Bayesian approach for quantifying and dealing with the uncertainty in parameter estimation (Box & Tiao, 1973; Berger, 1985) and model choice (King & Brooks, 2002; Spiegelhalter *et al.*, 2002), and (2) the development of Bayesian-geostatistical techniques (De Oliveira *et al.*, 1997; Gaudard *et al.*, 1999) for improved spatio-temporal modelling of egg distributions at sea. The case study under consideration is that of the North East Atlantic mackerel, for which an extensive dataset exists, but some findings and methods developed should have general applicability to other stocks assessed with the help of egg production methods.

1.1 Problem statement

Biomass estimates from egg surveys are an important part of the assessment procedures for many valuable stocks (Hunter & Lo, 1993). Management decisions on the European mackerel stock are based on an Integrated Catch at Age (ICA) model (Patterson & Melvin, 1996) "tuned" every three years by the egg survey-based SSB estimate. A series of problems and limitations regarding this estimate have been identified over the years (ICES, 1999a). Egg surveys for mackerel and horse mackerel carried out by ICES (International Council for the Exploration of the Sea) have gone a long way since their start in 1977, and continuous improvements have been made in the light of previous years experience. But the estimation methodology has not been making use of this accumulated knowledge, even in situations where a year's parameter estimate is suspected to be biased or inaccurate. The estimation procedure of the AEPM has no explicit protocol for "learning" from experience. Moreover, incomplete spatial coverage, together with the very coarse temporal scale used in the present method, in the order of a month or so, both could pose serious limitations to the accuracy and precision of the final estimate.

Finally, a new management procedure developed by other researchers carried out within the same research project will feed from the Bayesian stock biomass estimate, allowing novel analyses of a range of possible management scenarios for the Mackerel stock.

1.2 Purpose and description of the research

The present study investigates the use of Bayesian and geostatistical methods for the improvement of the fisheries management advice based on pelagic egg surveys. The project's objectives cover both practical and theoretical grounds. The development of a Bayesian-Geostatistical methodology, although not completely new, would be a substantial contribution to the research work being carried out in this direction by other researchers (for example Ribeiro & Diggle, 2000). No similar methodology has been yet applied to fisheries problems. The AEP method has been widely used in different regions and for different species (See Table 1.1), but no comprehensive review has been published of its limitations and problems, nor have the errors in its calculation been quantified.

The output of this work should have direct applicability on the pelagic stocks whose data are being used (Western Europe Mackerel and Horse mackerel) but also general interest for any other stock assessed with Egg Production methods. The objectives of my work can be summarised as follows:

Review the present methods of estimation of fish Spawning Stock Biomass of determinate spawners based on planktonic egg surveys (AEPM)

The traditional implementation of the AEP method will be reviewed, with special emphasis on the methods applied to the western European Mackerel and Horse mackerel stocks, as summarised in ICES (1999a).

Identify and quantify the sources of uncertainty in the Annual Egg Production Method

Information gathered on the parameters and assumptions present in the model will be used to quantify their individual contribution to the total error and bias in the egg production estimate. Data regularly collected during egg surveys are discarded or not used to their full potential. For example, fecundity information is summarised on a single mean estimate, despite the individual values being available, making variance quantification possible. Together with this, data have been gathered from various published and unpublished sources that relate to specific parts of the egg surveys, such as net performance (Anon., 1999) and fecundity estimation (Witthames, 2000). Some of the parameters analysed (fecundity and sex ratio, among others) will also be used when calculating stock biomass from egg abundance estimated by the geostatistical methods.

Develop a bias and error model for the Annual Egg Production Method, using Bayesian analysis to convey the uncertainty around estimated parameters and from inherent assumptions

The analysis of the sources and magnitude of uncertainty will allow for the construction of a probability distribution that would summarise all of that information. The probability distribution will be used as a constant of proportionality linking the standard biomass

or egg abundance estimates to a posterior probability distribution of values around the mean.

The implementation of this analysis is already being carried out as a set of computer algorithms that will improve dramatically computational speed, thus providing a fast and reliable platform for simulation studies and subsequent analyses.

ANALYSE WITHIN THE BAYESIAN PARADIGM A GEOSTATISTICAL SPATIO-TEMPORAL ESTIMATOR OF PLANKTONIC EGG ABUNDANCE

Researchers in the Environmental & Mining Engineering Research Group at Imperial College, have developed a geostatistical estimator of pelagic egg abundance (EC Research Project GBMAF, 1999/1253). Their aim has been to improve the present methodology by using more sophisticated and precise techniques for dealing with spatial data, and by making use of the whole time series when modelling the egg abundance surface over an area. Building on that work, a Bayesian framework will be developed to accommodate the geostatistical methods, thus providing a more precise and realistic estimate of egg abundance and fish biomass.

APPLY THESE METHODS TO THE WESTERN EUROPEAN STOCK OF MACKEREL AND HORSE MACKEREL

The ICES dataset for Mackerel and Horse mackerel egg abundances will be used in the development and testing of the methods. A final implementation specific to the characteristics of the North East Atlantic mackerel stock will be formulated for use by the ICES Working Group.

EVALUATE THE POSSIBLE USE OF THE AVAILABLE METHODS, BAYESIAN AND NON-BAYESIAN, ON THE MENTIONED STOCKS, AND IDENTIFY THE ADVANTAGES AND PROBLEMS ASSOCIATED WITH EACH OF THEM

Finally, the whole set of methods will be tested and evaluated through simulation procedures. Traditional AEPM, Bayesian AEPM, Geostatistical and Bayesian-Geostatistical approaches, together with the Generalized Additive Model (GAM) estimator(Augustin *et al.*, 1998), will all be compared in performance by means of simulated egg abundance datasets. The adequacy of present survey protocols to each of the methods, and possible improvements within the present sampling effort constraints will also be evaluated.

1.3 Ichthyoplankton and fish stock biomass estimation

The idea that the biomass of a fish stock can be quantified from its spawn is not a novel one (Hensen & Apstein, 1897), but its practical application in stock assessment is relatively recent. The original formulation of the most commonly applied of egg production methods, the AEPM, dates back to 1964 (Saville, 1964). Earlier studies concerned with ichthyoplankton abundance and spatio-temporal dynamics do not appear to attempt to use that information for inference on the abundance of adult fish.

This was until the 1980's the only available procedure for estimating SSB from ichthyoplankton surveys (Hunter & Lo, 1993). Other methods exist that apply to species with different biology, like the Daily Egg Production Method (DEPM), or are based in different procedures, like the Daily Fecundity Reduction Method (DFRM). The DEPM is in common use worldwide for the assessment of a number of stocks of small pelagics, mostly sardines and anchovies.

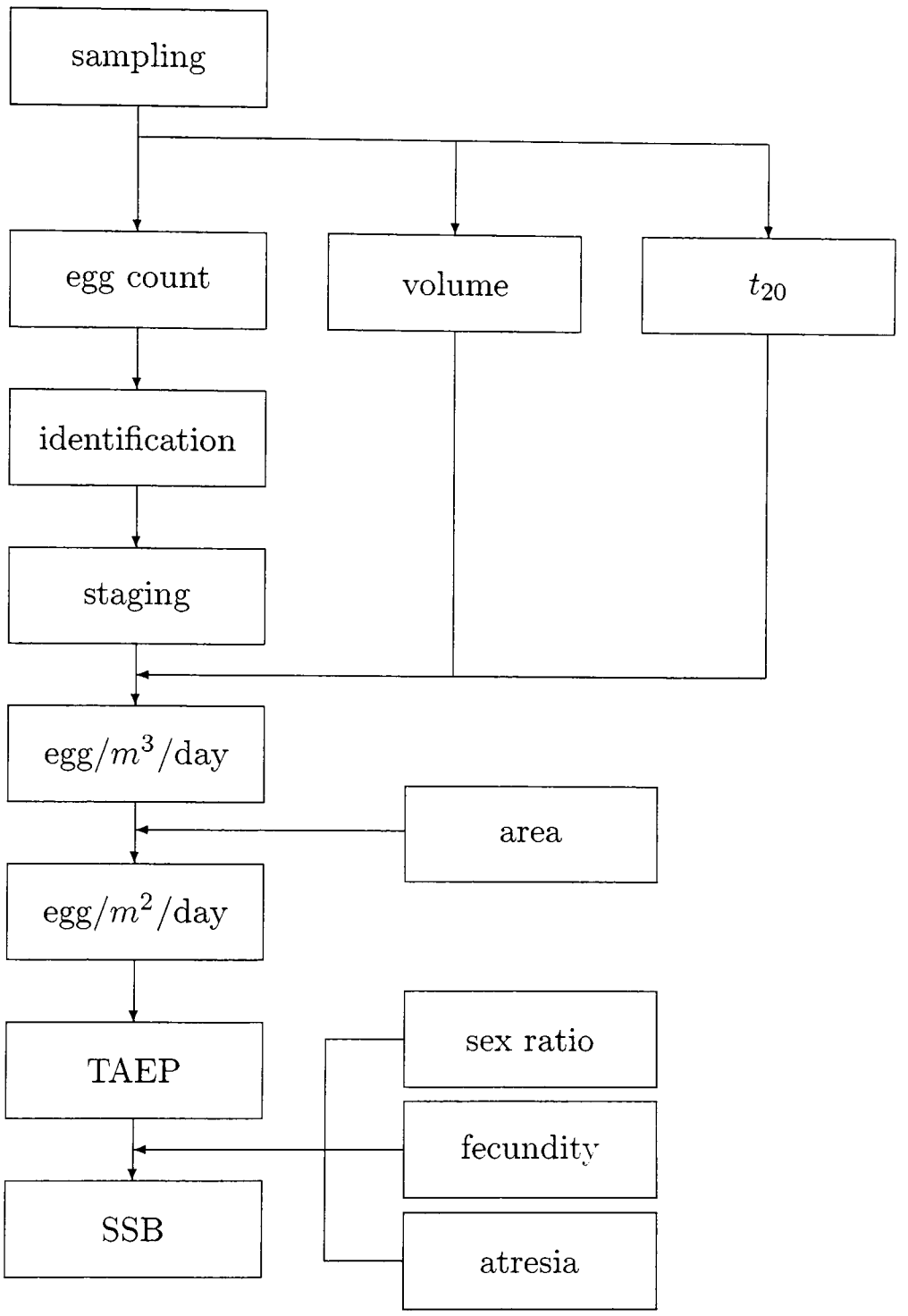


Figure 1.1: Flow diagram of the process behind the Annual Egg Production Method. All steps are explained in the text. t_{20} is the temperature at 20m.

The number of stocks where egg production methods have been applied is quite restricted, mostly due to the high costs involved. Apart from the cost of running the surveys themselves, the application of egg surveys to a new species depends on detailed studies of its reproductive biology, egg development dynamics, and egg distribution in time and

space. A total of nine studies have been found that apply the AEPM for estimating stock biomass (Table 1.1), although in some cases a single exploratory survey was carried out.

1.4 The Bayesian paradigm: modelling uncertainty

The recent upsurge in the use of the Bayesian approach in fisheries stock assessment could be ascribed to the confluence of two different developments. On the one hand, increasing computing power and developments in numerical computation have brought within reach of most fisheries biologists a set of methods previously unavailable. Second, a growing dissatisfaction with some of the outcomes of standard stock analysis (Hilborn, 2003) has prompted renewed interest in a paradigm that incorporates uncertainty in a most coherent way (Punt & Hilborn, 1997).

Uncertainty in fisheries biology is a common presence: sampling fish populations is a complex process, some parameters in stock models are notoriously difficult to estimate (e.g. natural mortality), and natural populations are under the influence of many processes, some of which escape our understanding. If we agree with Hilborn & Walters (1992) in their view that the main objective of fish stock assessment is to quantify the effect of alternative management options, it follows naturally that any advice should incorporate some measurement of how precise it is or how much it can be relied upon.

Leaving aside the philosophical and mathematical debates on the strength of the foundations of either frequentist or Bayesian statistics, a number of advantages have been mentioned in the literature in favour of Bayesian thinking (Gelman *et al.*, 1995), among them that,

1. Facilitates common-sense interpretations of statistical conclusions
2. Permits meaningful assessments in confidence regions
3. Enables the incorporation of prior knowledge into the analysis
4. Leads to more accurate estimators when prior prior knowledge is available and accurate
5. Obeys the Likelihood and Sufficiency principles

When confronting fisheries issues, the principal arguments in favour of the Bayesian paradigm could be summarised in two: it fulfils the need for estimating proper probabilities of alternative hypothesis in order to evaluate alternative management actions, and permits the incorporation of previously gathered knowledge for better estimating the full uncertainty in our understanding of nature (Punt & Hilborn, 1997). The conceptual simplicity and elegance of the Bayesian perspective, compared with the growing complexity with which some of the stock assessment models presently in use attempt to deal with uncertainty, must also be noted.

A complete presentation of Bayesian thinking can be found in Smith & Bernardo (2000) and Berger (1985). More practical texts are those of Box & Tiao (1973), Gelman *et al.* (1995), Lee (1997) and Congdon (2001).

Table 1.1: Summary of published studies presenting the results of applications of the AEPM for stock biomass estimation. Uncertainty in the reproductive biology (determinate/indeterminate spawning) is indicated by an interrogation mark.

Reference	Species		Location	Duration	Use
Sundby & Solemdal, 1984	cod	*Gadus morhua*	Lofoten, Norway	1983	Research
Lockwood, 1988	mackerel	*Scomber scoumbrus*	NE Atlantic	1977-2001[a]	Absolute SSB[b]
Alheit, 1988	sprat	*Sprattus sprattus*	North Sea		
Eaton, 1989	horse mackerel	*Trachurus trachurus*	NE Atlantic	1989-2001[a]	Absolute SSB
Kim & Gunderson, 1989	walleye pollock	*Theragra chalcogramma*	Alaska	1981	Compare to acoustic & trawl surveys
Picquelle & Megrey, 1993	walleye pollock	*Theragra chalcogramma*	Alaska		
Grégoire *et al.*, 1997	mackerel	*Scomber scoumbrus*	Eastern Canada	1983-96	Absolute SSB[c]
Bulman *et al.*, 1999	blue grenadier	*Macruronus novaezelandiae*	Tasmania	1994-95	
Fox *et al.*, 2000	cod	*Gadus morhua*	Irish Sea	1995	Comparison to VPA
	plaice	*Pleuronectes platessa*	Irish Sea	1995	Comparison to VPA
	sole	*Solea solea*	Irish Sea	1995	Comparison to VPA
Armstrong *et al.*, 2001	cod	*Gadus morhua*	Irish Sea	1995	Comparison to VPA
	plaice	*Pleuronectes platessa*	Irish Sea	1995	Comparison to VPA
	sole	*Solea solea*	Irish Sea	1995	Comparison to VPA

[a]Every three years
[b]Considered relative from 1998
[c]Normal distribution for AEPM curve assumed

Table 1.1: Continued

Reference	Species	Spawning mode	Days[d]	Stations	Sampler	CV (SSB)	Z	Stages[e]
Sundby & Solemdal, 1984	cod	determinate	15/82	415	Bongo (?)			I-VI
Lockwood, 1988	mackerel	determinate (?)	89/182 [f]	1440	Bongo/CalvNet	0.4		I
Alheit, 1988	sprat							
Eaton, 1989	horse mackerel	indeterminate (?)[g]	89/182	1440	Bongo/CalvNet	0.33		I
Kim & Gunderson, 1989	walleye pollock	determinate (?)	28/90	281	Bongo	0.09-0.43		I-II
Picquelle & Megrey, 1993	walleye pollock							
Grégoire *et al.*, 1997	mackerel	determinate (?)	12/170	65	Bongo	0.25 [h]		I,V
Bulman *et al.*, 1999	blue grenadier	determinate	34/93	238	CalvNet	0.14	0.5[i]	I-VIII
Fox *et al.*, 2000	cod	determinate (?)	83	1013	Bongo	15-21	0.25	I-V
	plaice	determinate	83	1013	Bongo	17-36	0.19	I-V
	sole	determinate (?)	83	1013	Bongo	14	0.39	I-V
Armstrong *et al.*, 2001	cod	determinate (?)	83	1013	Bongo	15-21	0.25	I-V
	plaice	determinate	83	1013	Bongo	17-36	0.19	I-V
	sole	determinate (?)	83	1013	Bongo	14	0.39	I-V

[d]Days of sampling/length of the spawning season
[e]Egg development stages incorporated in the AEP estimate
[f]For the 1998 survey
[g]Unpublished report. HOMSIR EU project
[h]For the egg production estimate
[i]Value assumed, actual estimate was zero

1.4.1 Bayesian methods in fisheries assessment

The first suggestion on the suitability of Bayesian analysis for comparing alternative fisheries policies can probably be found in the work of Walters & Hilborn (1976), although the first comprehensive application might be that in Bergh & Butterworth (1987). Bayesian methods have rapidly become more common in the fisheries literature, and this tendency is still on the increase.

Applications range from wide studies of various parameters accros stocks and species. be those biological (Myers & Mertz, 1998; Michielsens & McAllister, 2004) or fleet-related (Harley & Myers, 2001), to complete assessment models, either simple (Hoenig *et al.*, 1994; Meyer & Millar, 1999; Millar & Meyer, 2000; McAllister *et al.*, 2004) or very complex (Breen *et al.*, 2003; Bull *et al.*, 2005). In all these uses, the Bayesian paradigm has provided a coherent, integrated, flexible way of using previous information in ata analysis and fully considering the amount of uncertainty present in data and model. An indication of the growing acceptance of Bayesian methods in fisheries assessment can be obtained by considering the increasing space devoted to them by a range of textbooks and general treatises on the subject Hilborn & Walters (1992); Quinn & Deriso (1999); Hart & Reynolds (2002); Walters & Martell (2004).

1.5 Geostatistical methods

Geostatistics refers to the study of phenomena that vary in space and/or time (Deutsch & Journel, 1998). The tools provided by geostatistics aim at explaining and modelling spatial and temporal variability by use of a basic paradigm: unsampled values of a random variable are characterised by a probability distributions than models uncertainty around those values. Other applications include prediction and experimental design (Cressie, 1991). Most of the information used for modelling this probability distribution comes from values obtained at nearby locations. A basic tenet of all spatial statistics is that data that are close together in space (and time) are often more similar. The objective of spatial models is to incorporate this dependence into the generating mechanism (Cressie, 1991). From this point of view, spatial models depart considerably from many other modelling approaches that assume observation error to be the only random mechanism at work.

Geostatistics is largely based on the concept of a random function (RF), whereby the set of unknown values is regarded as a set of spatially dependent random variables (Goovaerts, 1997). The possible realizations of the random variable at a given location represent the local uncertainty about its value. The set of realizations of the random function models the uncertainty about the spatial distribution of the attribute over the study area. A random function is thus defined as a set of usually dependent random variables defined over some field of interest that can take a variety of outcome values according to some probability distribution function.

An important concept to consider when dealing with random functions is that of stationarity. As repeated samples are not commonly available at the same location, an

assumption is introduced that the phenomenon under study is spatially "homogeneous" within the study area. This means that the RF model is also assumed to be stationary within the area considered. This assumption allows the use of a number of measurements separated by the same distance, although in different locations, for inference of the relationship between the random variables separated by such distance. Stationarity is a property of the RF model needed for inference, and it is not a hypothesis that can be tested (Goovaerts, 1997). However, this strong assumption can be relaxed, for example if estimating through Ordinary Kriging, to apply only within the search neighbourhood defined for prediction.

Geostatistics has evolved from this set of basic concepts into a wide range of methods and procedures for the study of spatial phenomena. Although originally developed for the study and estimation of ore deposits (Journel & Huijbregts, 1978), it is now in use across a number of disciplines, such as meteorology (Berke, 1998), hydrogeology (Kitanidis, 1997), environmental chemistry (Kitanidis & Shen, 1996), pollution (Qian, 1997), and of course fisheries (Petitgas, 1993).

1.5.1 Use of geostatistics in fisheries assessment

Spatial processes in fisheries were studied from an early stage in the discipline (Beverton & Holt, 1957), but detailed consideration has always been hampered by the difficulty of collecting spatial information for fish distribution. The classical ecological consideration of fish populations as homogeneous, closed units, might have been a significant factor in the late development of the field of spatial fisheries analysis. Geostatistics has played an important role since fisheries biologists began looking carefully at spatial questions in abundance estimation (Booth, 2000). Fishery-independent surveys of fish abundance saw the first uses of geostatistical estimators for the analysis of trawl (Petitgas & Poulard, 1989), longline (Pelletier & Parma, 1994) and egg surveys (Petitgas, 1997) of fish or crustaceans (Freire *et al.*, 1991; Maynou & Sardá, 1998), and most notably acoustic surveys for a variety of species (Guillard *et al.*, 1992; Petitgas, 1993).

Early interest in geostatistics in fisheries research concentrated in its use for stock assessment (Conan, 1985), and survey design (Gohin, 1985). This second field has seen intense use of geostatistics in recent years, as the use of spatial covariance enables the use of a variety of non-random survey designs (Petitgas, 2001). Fisheries data, including scientific surveys, often departs heavily from assumption of random sampling, as samples are not completely independent from one another in space or time, and the sampled population is often spatially structured.

A particular field where geostatistical methods have been applied intensively is in the analysis and estimation of acoustic surveys (Petitgas, 1993). Sonar has a wide use as a tool for observing the behaviour and estimating the extension of concentrations of marine fish (Gunderson, 1993). In contrast to other methods, it is almost non-intrusive and offers

the potential for almost continuous sampling along a series of transect lines. But this deployment strategy makes it depart significantly from the limitations of random sampling theory. First, samples taken along the same transect are serially correlated and, second, transects are rarely random and tend to cover instead most of the area of interest. Geostatistics provides a modelling approach for properly analysing such data. For example, Maravelias *et al.* (1996) used robust and classical variograms, simple and nested models, and point kriging in their analysis of herring acoustic surveys off the eastern Scottish coast. They applied a robust variogram estimator of Cressie & Hawkins (1980) due to the presence of a number of extremely high values that could not be regarded as outliers, and which tend to distort the real variogram patterns (Goovaerts, 1997).

Following on the initial work of Conan (1985), a number of authors have applied geostatistical methods to shellfish stocks. The comparatively limited mobility of most crustaceans render then amenable to simple geostatistical methods. Abundance data for populations of a range of taxa have been analysed with a variety of kriging techniques. The spatial structure of the populations of different decapods Freire *et al.* (1991), epibenthic galatheids (Freire *et al.*, 1992), brachyurans (Gurriarán, 1993), and Norway lobster (Fariña *et al.*, 1994) surveyed along the coast of Galiza (NW Spain), were all analysed using semivariograms and point kriging. Stocks of *Nephrops norvegicus* in the northwestern Mediterranean were also analysed through point kriging (Maynou & Sardá, 1998). What all these studies have in common is that the main interest was not in estimating the total biomass of the resource over the sampled area, but in investigating the spatial pattern of relative abundance, and how they related to the survey design in place.

Geostatistics ave been applied to ichthyoplankton surveys, including the North East Atlantic (NEA) mackerel dataset. Fletcher & Sumner (1999) present an application to eggs and larvae of sardine (*Sardinops sagax*) off western Australia. Again a dual objective directed this study: estimation of abundance and analysis or survey design. Correlograms and variograms are estimated from a limited number (50) of plankton tows, collected at varying intervening distances. Hauls were replicated to collect information on small scale and sample variability. The small sample size prevented kriging being carried out based on the directional variograms, that appeared to fit better to the observed abundances. Patch size was then estimated, providing an useful indication of the preferred distance between sampling stations. This was shown to vary with egg age, as patches get larger and less dense as eggs disperse. Unfortunately, the increase in precision achieved by simply adapting sampling intensity to patch size was found to be uneconomical. Alternatively, they proposed an adaptive sampling protocol that could then be analysed using geostatistics.

A series of studies have applied geostatistical methods to the NEA mackerel egg surveys dataset (Bez *et al.*, 1997; Bez & Rivoirard, 2000, 2001: Bez, 2002). A review of some of this work is included in Rivoirard *et al.* (2000). The authors explore a series of geostatistical methods for the analysis and estimation of egg abundances, with an emphasis on the use of widely available covariates, such as Sea Surface Temperature (SST). The

first of these studies (Bez *et al.*, 1997) presents estimates of the centre of gravity (average position of the abundance) and inertia (statistical dispersion of the samples) for the 1983 to 1995 surveys. Analyses are based not on the original sample locations but on a regular grid where each sample is assigned to the centre of a 0.5° longitude by 0.5° latitude square. Changes in the centre of gravity are detected and cannot be explained by changes in sampling intensity alone, giving an indication of the displacement of mackerel spawning to the west, as noted more recently by other analyses (Beare & Reid, 2002).

The relationship between egg and larvae abundance and SST was then investigated through variography (Bez & Rivoirard, 2000). The individual-based approach advocated by these authors was based on the probabilistic modelling of the relationship between each sample and the value of SST at the same location. To investigate whether the apparent correlation between SST and ichthyoplankton abundance is the result of a causal relationship, egg and larval abundances were relocated along the north-south and east-west axes and new inertiograms were then estimated. If a relationship existed between a population (mackerel ichthyoplankton) and a field with homogeneous variations (SST), such translations would lead to an increase in inertia. No such increase was found for most egg stages and larval ages, although doubts were cast on the assumption of homogeneity of the temperature field. Additionally, when estimating abundances over the whole area, individual based statistics will count each sample proportionally to the number of individuals. The question of the lack of independence between individuals, eggs or larvae, collected in the same sample remains unanswered. The previous exercise was later extended to four surveys, 1986 to 1995 (Bez & Rivoirard, 2001). Consideration was made of the limitation that transitive geostatistics, as proposed in the previous publications, depend on the availability of detailed data for the covariate chosen, SST in this case. Such data might not always exist or might be difficult to acquire.

Covariograms were again applied to the ICES mackerel data in Bez (2002). The model presented allowed for different anisotropy parameters, direction and anisotropy coefficient, to be estimated for each of the functions forming the covariogram. Values for the nugget, ranges and sills were estimated by least squares, while the anisotropy parameters were manually set. The estimated variance in egg abundance appeared to be quite low ($CV = 8.4\%$), despite the fact that most of the spatial structure was concentrated at scales smaller than the inter-sample distance.

1.5.2 Bayesian geostatistics: developments and applications

In common with other areas of statistics, the last two decades has seen growing interest in the application of the Bayesian paradigm in geostatistical methods. Developments in this area are however very recent. For example, a search for *Bayesian* and *variogram* or *kriging* in the Web of Science database (ISI, July 2004), finds its first reference in 1987 (Omre, 1987).

Some of the seminal work focused on simple universal kriging where conjugate formulations could be appropriate (Omre, 1987; Omre & Halvorsen, 1989). In the first of these studies, an extension is presented to the linear Bayesian model to a spatial setting. Simple kriging is extended to include uncertainty in the parameters of the random function. Prior distributions, termed "qualified guesses" by the author, are limited to the normal family, but no formal presentation is made of either priors or likelihood function. The exercise was later completed by incorporating universal kriging (Omre & Halvorsen, 1989). Prior information is applied in both cases to the mean function only, not accounting for uncertainty in the covariance structure.

Another initial step can be found in the work of Handcock & Stein (1993). The Matérn function is applied to the modelling of elevation data, and a comparison is made between Bayesian, maximum likelihood and standard variogram fitting. Use is made of uniform and normal priors, but choices are subject to sensitivity analyses. Other studies that investigated the application of the Bayesian approach to the estimation and prediction of Gaussian random fields are Brown *et al.* (1994), Handcock & Wallis (1994), and Christakos & Li (1998).

Emphasis is made in some studies on the usefulness of formally incorporating accumulated knowledge about the modelled process into the estimation procedure. Estimation of covariance parameters in data-poor situations can benefit from the experience gained through analyses of similar situations (Pardo Iguzquiza, 1999), and the Bayesian approach allows, through the use of prior probability distributions, for previous results to be incorporated in the analysis (Gelman *et al.*, 1995). Bayesian methods have also been applied to the problem of variogram model selection, either through the use of a more general formulation that includes different functional models (Ribeiro & Diggle, 2000), or through Bayesian Model Averaging (BMA) (Ming Ye *et al.*, 2004). Only very tentative uses of Reversible Jump MCMC (Brooks *et al.*, 2003) have been found in geostatistics for model selection (Chaudhuri *et al.*, 2004).

Non-Gaussian fields can be modelled by assuming that the random field of interest is the result of an unknown linear transformation of a Gaussian random field. A function can then be chosen, such as the Box-Cox family of power transformations(Box & Cox, 1964), that allow for the best transformation parameters to be estimated from the data (De Oliveira *et al.*, 1997). The Bayesian approach applied by De Oliveira *et al.* (1997) computes the uncertainty introduced by the estimation of the transformation parameter together with the variogram and sampling error. This is shown to generally increase the variance of the estimated values. Use of an incorrect transformation, selected due to the effect of outliers, could significantly bias the results obtained.

Finally mention must be made of the line of work of Diggle and co-workers, who have been applying Bayesian analysis to a wide range of spatial problems. An extensive coverage of the ideas and their implementation can be found in Diggle *et al.* (1998). Geostatistical problems are treated here as deriving from generalized linear models (GLMs).

Observations at sample locations x_i form a GLM conditional on an unobserved stochastic spatial process $S(x_i)$,

$$Y_i|S(x_i) \sim N\{\mu + S(x_i), \theta^2\} \tag{1.1}$$

The ideas exposed in this paper have also had methodological development with the release of a software library in the *R* language, *geoR*, that provides functions for 2D kriging and MCMC algorithms for computation of variogram and kriging parameters (Ribeiro & Diggle, 2000).

Problems of estimation of a total value over a surface or volume, like the one encountered in biomass estimation from egg surveys, requires the use of block kriging. In this method, the average value of an attribute over a block of specific dimensions is estimated (Goovaerts, 1997), instead of a point estimate at a number of locations as is done in other kriging types. Bayesian applications of block kriging are not common. Todini (2001) presents a Bayesian analysis of rainfall data, where precipitation fields estimated from meteorological radars are conditioned on point measurements from a network of rain-gauges. The main objective was to ensure radar-based estimates over a whole area agreed with the data collected at ground level, but uncertainty in the later measurements was incorporated in the analysis. The actual implementation relied on a Kalman-filter, a common method for dealing with Bayesian linear models but that has proved both inflexible and cumbersome (Meyer & Millar, 1999).

The number of developed applications of Bayesian geostatistics is still fairly limited, as previously presented. Although progress is fast, many problems remain unsolved, limiting the range of uses and applications. No example could be found of its use in fisheries or population ecology, where uncertainties due to sampling and structural error are often larger than in other fields.

1.6 The Northeast Atlantic mackerel fishery

Until relatively recently, mackerel has not featured among the most important European fisheries, neither in terms of catch nor market value. It was the important changes brought in the mid-1970s by the declaration of the 200-mile Economic Exclusive Zone (EEZ) by the countries commonly visited by the UK distant-water cod fleets that turned the interest of the industry towards other species, including mackerel. Only four years after the first two British distant water trawlers started fishing for mackerel, in 1978, this species' landings (187,000t) exceeded those of cod for the first time. Growth continued when the herring (*Clupea arengus*) fishery closed in the mid-1970s and the herring fleet moved to mackerel as its target species.

In the years since 1969, the fishery for Northeast Atlantic mackerel has oscillated around an annual catch of 700,000t (Figure 1.2). Catch in recent years has exceeded the recommended Total Allowable Catch (TAC). While the total quota was set at 683.365t for 2002, catches reached 717,882t. More than 10 countries have reported catches in 2002.

The main mackerel catching countries in recent years have been Norway, Scotland, Ireland, Russia, Netherlands and Spain. Significant catches are taken in the North Sea (over 360,000t in 2002), although they are considered to be part of the Western stock, and only 10,000t of those are consigned as North Sea stock catch.

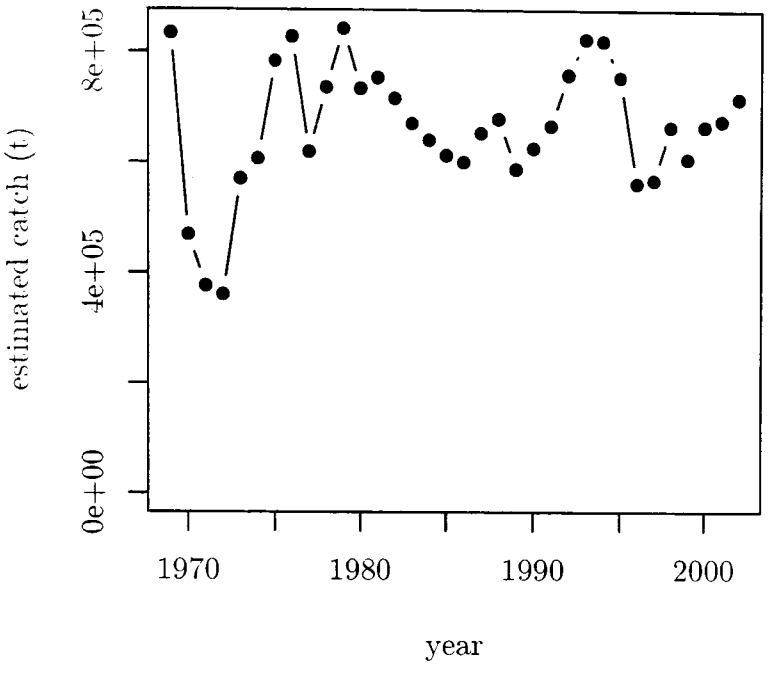


Figure 1.2: Catches (landings + discards) of Northeast Atlantic mackerel between 1969 and 2002 (ICES, 2004b).

Fishing mortality is distributed among a range of ages 1.3). Relative importance for the fishery of different age groups is affected by strong year classes, a not unusual occurrence in mackerel, in common with other small pelagics. The lower fishing pressure exerted on the younger ages appears to be related to the higher prices the market places on larger fish (Lockwood, 1988).

For further general information, a comprehensive coverage of the Atlantic mackerel fisheries, biology and research, can be found in Lockwood (1988). An account of the earliest studies and descriptions of the pre-industrial fisheries can be found in Cornish (1883).

1.6.1 Stock structure

The exact stock structure of the Northeast Atlantic mackerel is still an issue under debate, but it is presently assumed as a single unit with three components (ICES, 2004b). The western component is the largest, considered rather arbitrarily to extend from the French-Spanish border to the northern Scottish islands (Figure 1.5). Mackerel found in the waters around the Iberian Peninsula form the southern component. The division between these two sub-stocks appears to have no biological basis (Uriarte & Lucio, 2001), and remains in place probably only due to management reasons. Finally, the North Sea component, still assessed and managed slightly independently from the other two, used to be of major

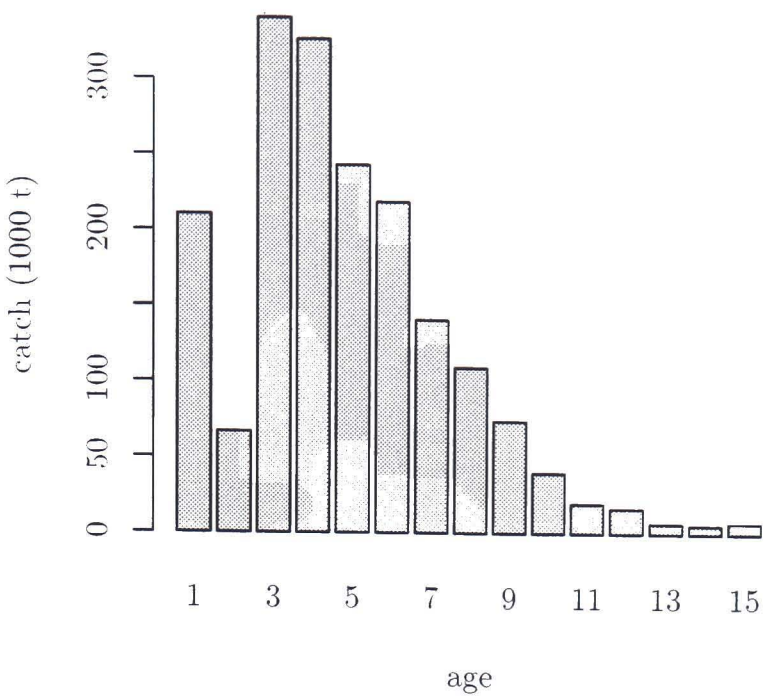


Figure 1.3: Catch-at-age of Northeast Atlantic mackerel in 2002 (ICES. 2004b).

importance, but supports a limited fishery and a much reduced biomass. Stock size was estimated at 3.5 million tons in the 1960s, while is considered to be now just over 200,000t.

The work developed in this study concentrates on the western, and to a minor part the southern, components. The western section has been assessed through egg surveys since 1977, while the southern section has been fully assessed using egg surveys only since 1995. Egg surveys are also carried out for the North Sea stock, albeit with lower frequency and intensity.

1.6.2 The European mackerel egg surveys

Since 1977 scientists from a number of countries, and under the auspices of ICES (International Council for the Exploration of the Sea), have been conducting triannual ichthyoplankton surveys directed at mackerel (*Trachurus trachurus*), and to a minor extent at horse mackerel (*Scomber scombrus*). These surveys have become an integral part of the assessment process for these stocks, and provide estimates of Spawning Stock Biomass for *tuning* of the abundance indices derived from catch and effort data. Estimates of SSB are obtained according to the Annual Egg Production Method (AEPM) (Alheit, 1988). Estimates of stock size from egg surveys in this area have historically been considered to be absolute, although uncertainties in some of the estimated parameters forced the relevant ICES Working Group in 1999 to incorporate them into the ICA-based assessment as relative abundance indices only (ICES, 1999b).

The initial surveys covered most but not all of the spawning area and season, in what was then considered the Western mackerel stock (Lockwood *et al.*, 1981). Three stock units formed at that time the basis for management: the Western stock. extending during the spawning season from the south of France to the north of Scotland: a Southern stock. spawning in Iberian waters; and the North Sea stock. The first two are now treated as a

single stock, although both components are assessed somehow separately. The North Sea stock is still treated as a separate entity in assessment terms. Fish from all three units can be found wintering off the Norwegian coast, where they form dense feeding aggregations apparently linked to water temperature (Reid *et al.*, 2001).

Surveys for North east Atlantic mackerel started in 1977 in the Western waters of the European continental shelf, from the Spanish-French border to the Northwest of Ireland (Figure 1.4). From the initial survey coverage of 438 samples and 133 days at sea, achieved with 2 boats, surveys have improved in both coverage and intensity (Table 1.6.2). In 1995 sampling of significant intensity took place for the first time in the waters around the Iberian Peninsula, and from that point these surveys have covered the whole of the known spawning area and season of Atlantic mackerel.

year	samples	positive samples	days	No. boats	No. countries
1977	438	287	133	2	2
1980	600	343	143	3	3
1983	860	681	120	6	5
1986	834	666	108	5	5
1989	930	766	109	7	6
1992	1173	725	89	8	7
1995	1041	684	149	9	7
1998	1080	557	186	8	7
2001	1409	789	124	8	7

Table 1.2: Summary information of surveys undertaken in the NE Atlantic for mackerel eggs between 1977 and 2001.

A detailed description of the procedure involved in the north east Atlantic mackerel egg surveys forms part of Chapter 2, but a simplified account is presented here. Eggs are first collected by means of Gulf III (in the western area) or vertical Bongo (in the southern area) plankton samplers. After species identification and separation according to developmental stages (Lockwood *et al.*, 1977), egg abundances are estimated from the number of eggs found in the sample, volume of water filtered, and length of the development stage at the water temperature encountered. The procedure is repeated at as many locations as possible during a number of sampling periods (usually six) along the spawning season. Egg production estimates are finally combined with estimates of adult female effective fecundity to arrive at an estimate of SSB for Atlantic mackerel.

The method presently in use for the spatial and temporal integration of individual egg abundance estimates works by assigning all sampling carried out within one of the set periods to the central day of that period. Spatial and temporal interpolation are separate procedures, the first by simple averaging on unsampled rectangles that are surrounded by sampled ones (ICES, 1996), the second by constructing an annual egg production

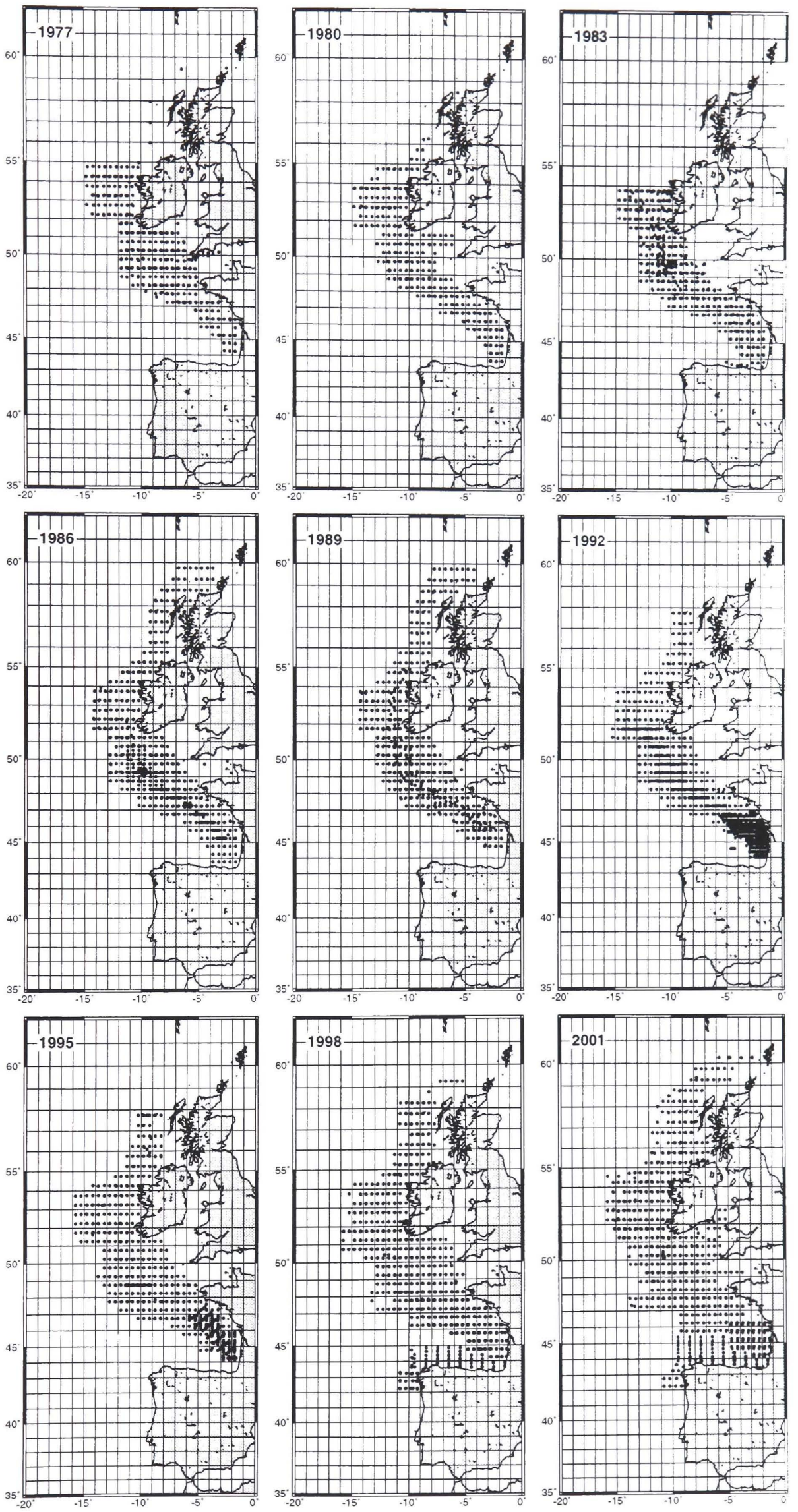


Figure 1.4: Locations sampled during the triannual ICES egg surveys.

curve based on each period estimates together with the assumed start and end dates (see Chapter 2 for a detailed description).

The survey design in place is strongly influenced by the usual constraints in ship time, but also by the estimation procedure outlined above. Much emphasis is placed on covering as much of the spawning area as possible during each period, and in sampling early and late enough in the season to detect the onset and close of the spawning period (ICES, 2000). Spatial resolution employed has been of half an "ICES statistical rectangle", meaning 30' latitude by 30' longitude. Although a standard survey area exists, sampling has been extended to the west over the years, and an adaptive sampling rule, where sampling stops only after a zero value is encountered, is now in place (Figure 1.5).

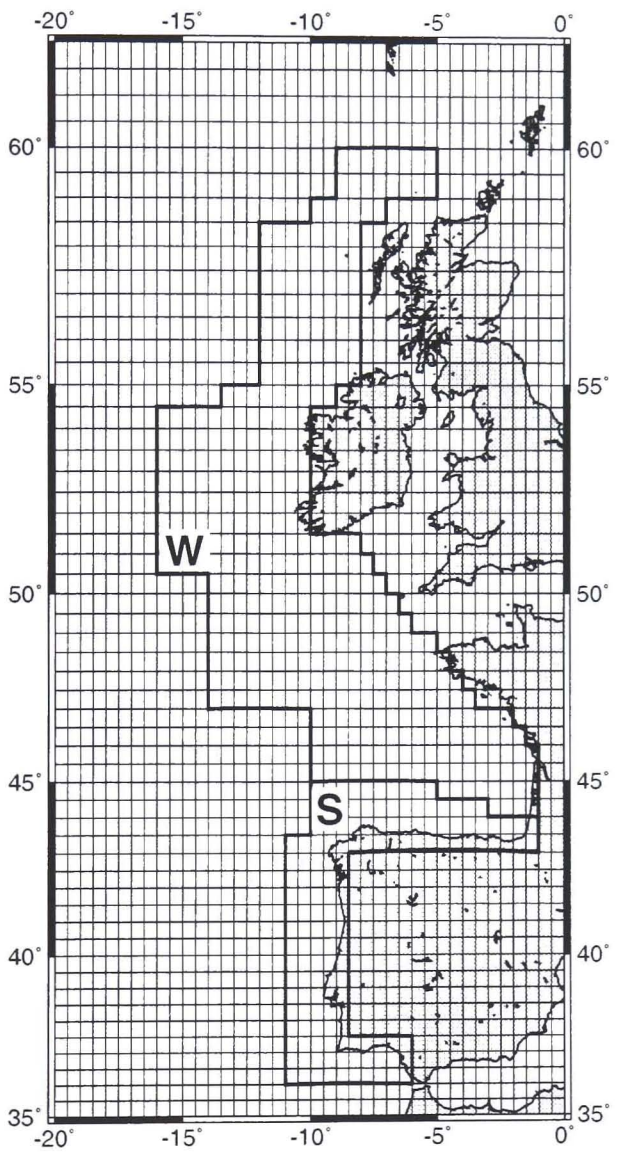


Figure 1.5: Standard survey areas considered in ICES mackerel egg surveys. W denotes the western component and S the southern one.

1.6.3 Exploratory analysis of the dataset

Before ambarking on an estimate of annual egg production and spawning stock biomass, much can be learned from an exploratory analysis of the ICES mackerel dataset. A graphical representation of the 3-dimensional distribution of samples allows us to see the survey

spatio-temporal design (Figure 1.6). Clearly observable in the first years, when sampling effort was a fraction of that deployed in recent surveys, is the intense spatial coverage achieved during relatively short time periods. The temporal distribution of sampling, concentrated around a number of short periods, can be clearly seen in 1980, where the increase in sampling was achieved by increasing the number of sampling periods. Also, the extension of sampling to the southern component can be observed in 1995 as the tail over the lower latitudes and dates.

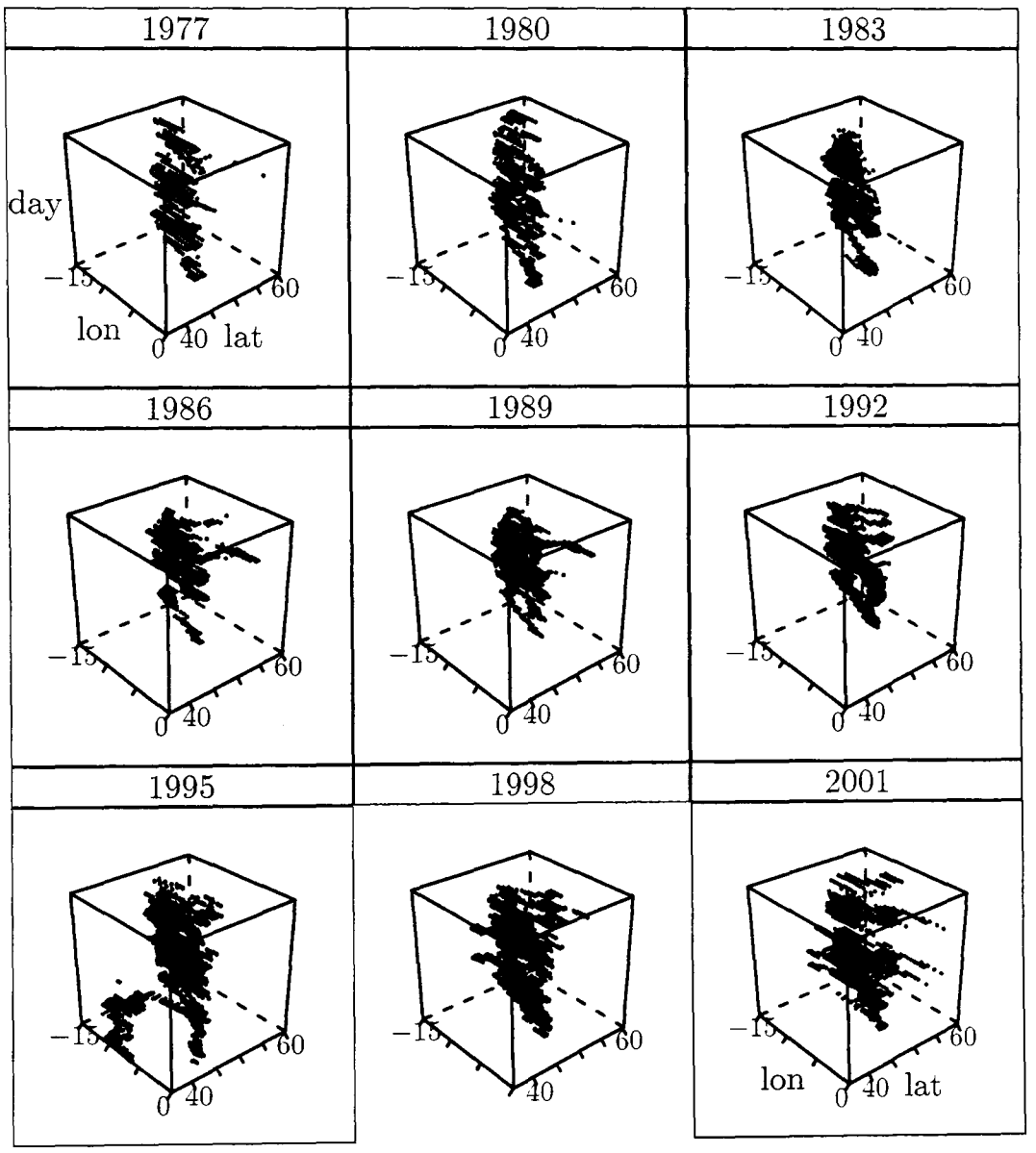


Figure 1.6: Three-dimensional representation of the spatio-temporal distribution of samples by year. Origin is set at at 40° N - 0° W.

Changes in the timing of spawning might be observed from plots of egg abundances by month (Figure 1.7, ICES NEA mackerel egg survey dataset), especially if shown, as is the case here, in comparison with the time distribution of sampling effort. Disparities in relative sampling effort and egg production indicate that the survey design, or its actual implementation, have failed to match each other as desired. For example, in 1992 the usual peak in sampling in June and July could only certify an early decline in spawning compared to previous years. Early spawning appears to be difficult to forecast, and as

extra sampling effort cannot be rapidly deployed even if significant spawning is detected at the onset of the surveys, the risk is always present of large numbers of eggs being sparsely sampled. That was the case in both 1992 and 1998, where the relatively low intensity sampling during the month of march encountered very high number of eggs.

Looking in more detail at spawning, patterns. are remarkably similar among years (Figure 1.8). Differences in scale and intensity appear more noticeable than differences in their start and end dates. There are often two peaks in spawning intensity and activity through the year. Spawning levels off 20 to 30 days before the final survey date in almost all years. The distribution of sampled egg abundance (as simple egg counts) at different dates, latitudes and longitudes, and for each year, can be observed in Figure 1.9. Again the temporal structure of sampling can be observed in the top panel, and how the increase in sampling effort, both in days at sea and boats, have made the limits between sampling periods less clear. The same can be said in the spatial scale (middle and bottom panels). Both the extension in coverage and intensity can be ascertained from these plots.

1.6.4 Assessment and management needs

The NEA mackerel fishery has been relatively stable over the last 20 years, both in terms of effort and catch, partly due to the sensitivity of the market to increase in landings. The largest fleet in the fishery, the Scottish fleet of pelagic trawlers, lands most of its catch in Norwegian ports, where better prices and higher demand are guaranteed. This market appears to drop in price as soon as supply surpasses demand. The fleet has thus generally restrained itself from landing large catches for purely meconomic reasons, and in consequence the annual Total Allowable Catch is usually not reached.

Despite this relatively comfortable situation, management of small pelagics tends to be complicated by the inherent variability of their populations (Spencer & Collie, 1997). Mackerel recruitment has also been found to vary with environmental factors such as turbulence and the North Atlantic Oscillation (NAO) (Borja *et al.*, 2002). This could increase the normal variability in stock size and recruitment found in this species, if environmental conditions where to become more extreme.

Under these circumstances, and considering the growing economic importance of mackerel to the European fleets, there is an increasing need to develop more precise and flexible assessment methods which could improve stock management by allowing much earlier detection of negative trends in population size and recruitment. Links might also be inferred between predictable environmental factors, such as NAO, and population size, distribution and catchability. Finally, a range of management options could be explored that proved resilient to the expected variability in population dynamics and environmental conditions.

A number of deficiencies in the present assessment procedure have been highlighted by the relevant ICES working group (ICES, 2004b). Regarding the egg surveys. the major concern expressed is that the assessment is over sensitive to the most recent surveys. Mackerel egg surveys in ICES have generally been considered as absolute estimates of

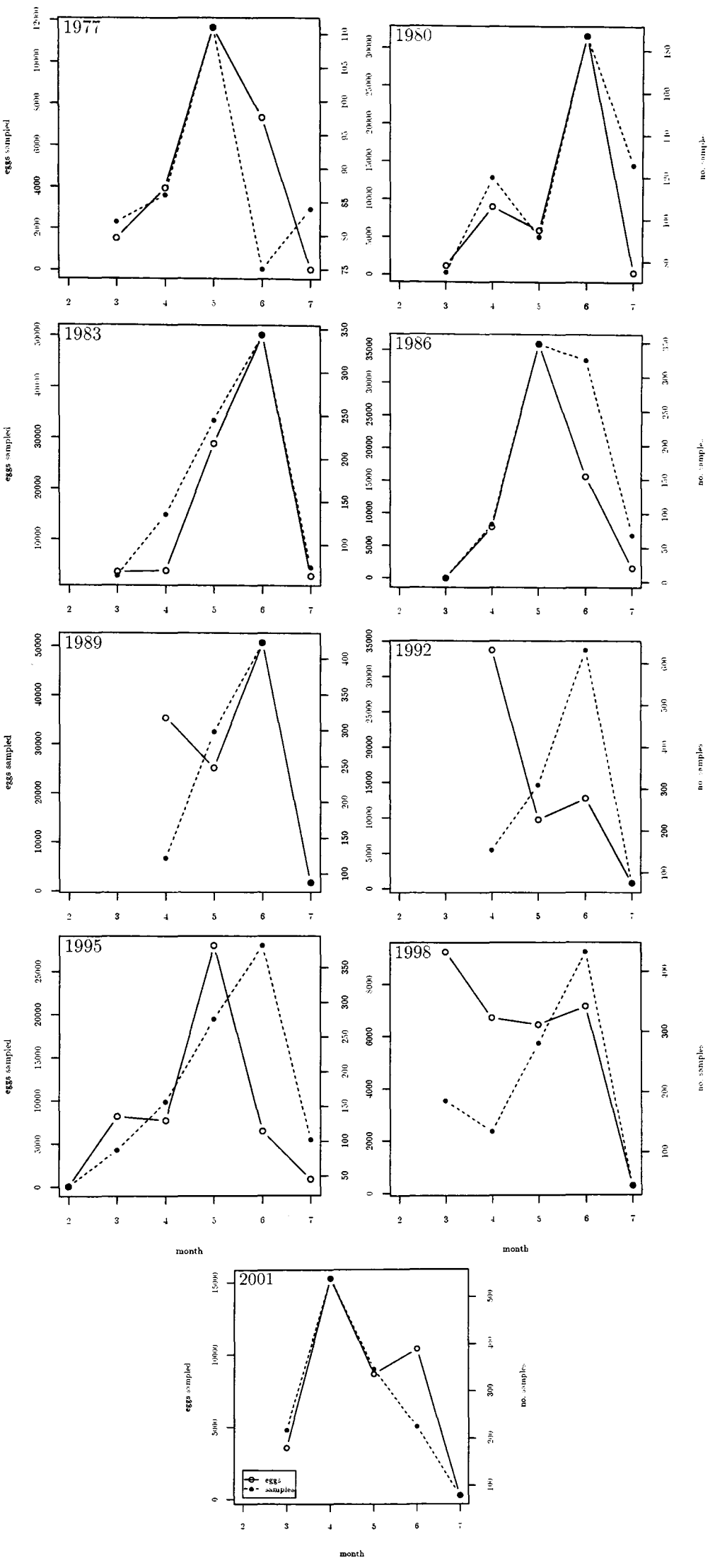


Figure 1.7: Egg counts (open circles) and sample sizes (dots) per month and year

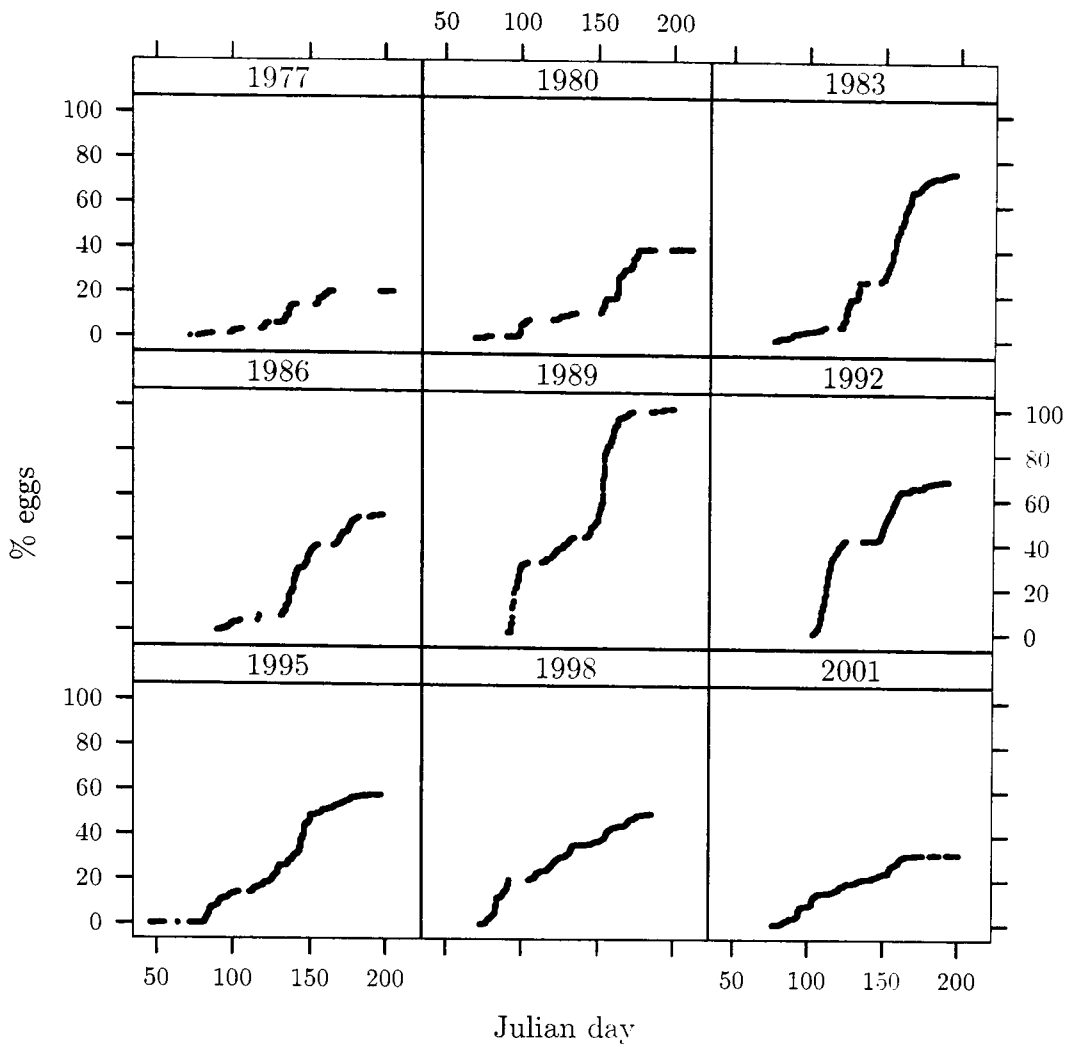


Figure 1.8: Time cumulative egg abundance by year. Abundances expressed as a percentage of the highest historical egg count.

SSB, although doubts have been expressed in some years about this assumption (ICES, 1999a). The assessment model currently applied to this stock, ICA (Patterson & Melvin, 1996), places considerable weight on the SSB estimates provided by the egg surveys, and so results are extremely dependent on the reliability of those estimates. Uncertainty in model parameters is presently not incorporated in the stock assessment. Comprehensive estimates of uncertainty in the SSB estimates derived from egg surveys should be a first step in addressing this issue.

1.7 Thesis outline

The present study attempts to quantitatively review the uncertainties present in the AEPM (Chapter 2). Then, and building on that knowledge, it explores the use of Bayesian geostatistical methods for the estimation of annual egg production. First, Bayesian spatio-temporal variograms are constructed that model the covariance structure of ICES mackerel egg survey dataset in the three dimensions (Chapter 3). These variogram models are then used in a Monte Carlo block kriging framework for the estimation of Total Annual Egg Production and its associated uncertainty (Chapter 4). Finally, both analyses are combined to assess the uncertainty present in the estimates of mackerel SSB from egg production surveys (Chapter 5). A comparison is also made in this last section with other

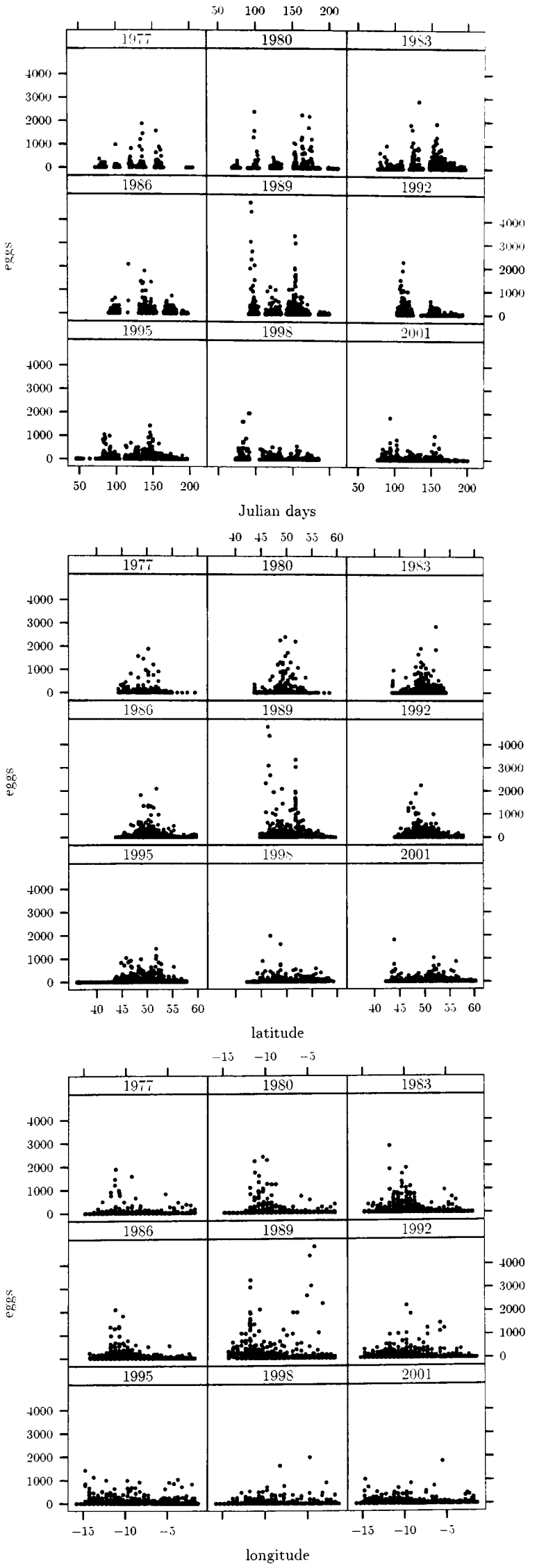


Figure 1.9: Egg abundances as raw egg counts vs. Julian day, latitude and longitude, and by year.

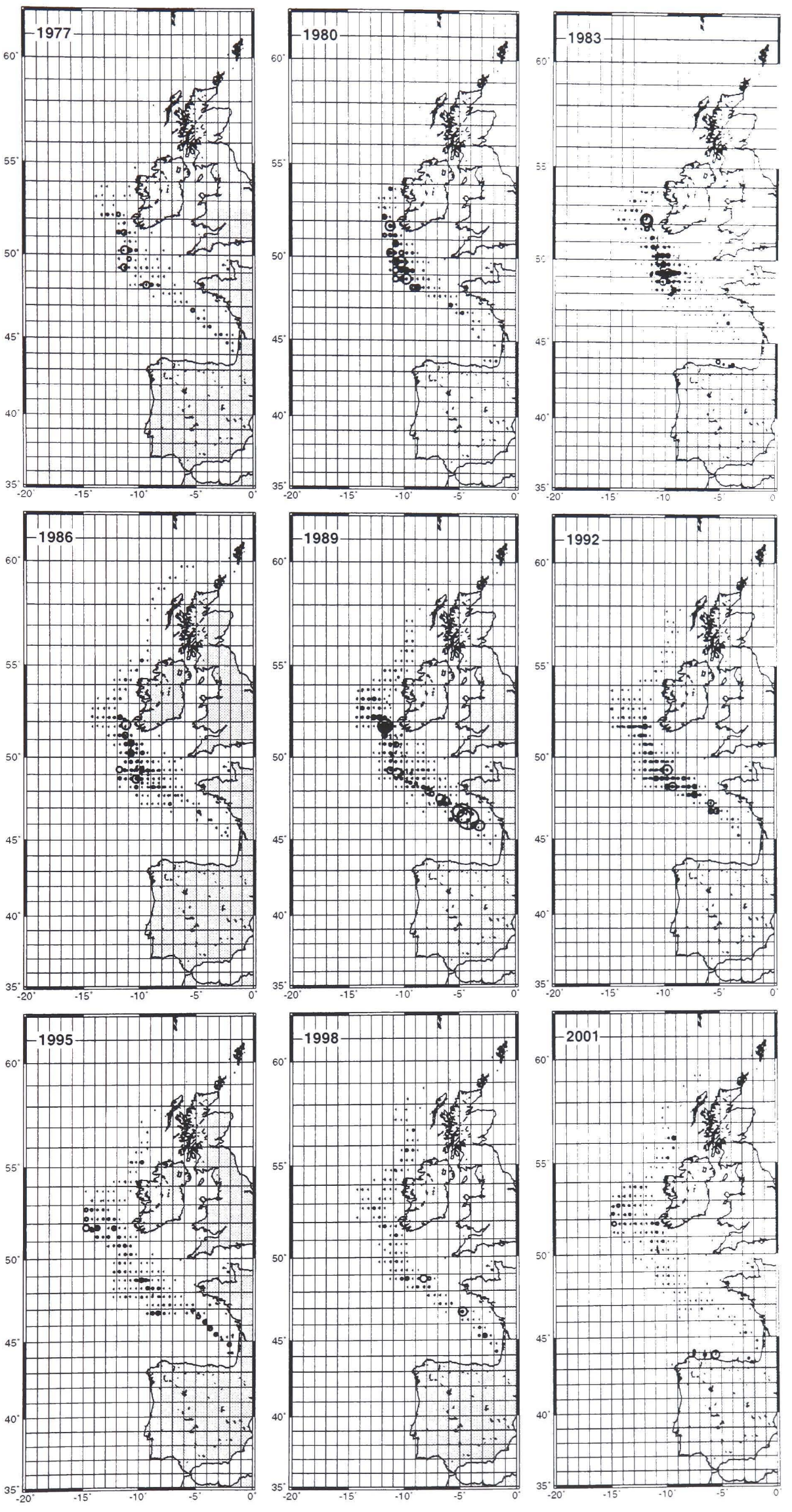


Figure 1.10: Abundance of mackerel eggs sampled during the triannual ICES egg surveys. Circles represent egg counts, not estimated egg abundances. Scale is common to all years.

methods, and a range of conclusions is presented on the usefulness and capabilities of these estimators.

Chapter 2

Review of the Annual Egg Production Method

Forty years after its invention (Saville, 1964, after Hunter & Lo, 1993), and with nearly 15 years of ICES mackerel egg surveys, the annual egg production method has provided a rigorous basis for fishery-independent stock assessment, albeit in a fairly limited number of situations (See Chapter 1). However, many of its implicit assumptions remain untested, and the possible sources of error and bias in its calculation are poorly understood.

Ichthyoplankton surveys for different egg production assessment methods follow a common sampling protocol. The ICES survey for mackerel and horse mackerel provides a representative example of the procedure. This survey, as it is carried out every three years in the waters of the Western Atlantic European facade, will be reviewed and analysed. Results and observations could be considered applicable to other assessments carried under the same framework.

2.1 Data collection

Ichthyoplankton samples are collected using some modification of the sampling nets used for other plankton work. High speed samplers consist of a solid frame to which a fine mesh net (between 250 and 500 μm aperture) is attached. A nose cone covers the mouth of the net. Two flowmeters are also attached to the frame, one of the inside and one on the outside of the net mouth. A metal weight keeps the sampler in the right angle. This must be almost horizontal for samplers towed in a near vertical V-shaped trajectory, such as the Gulf III sampler, or upright in the case of the CalvNet and other vertical *bongo* samplers. The sampler is either towed or lowered at a constant speed, either to the bottom or to a previously set maximum depth. The depth reached by the sampler is monitored by a sensor attached to it. When recovered, the net is washed and the sample is extracted. Flowmeters are read to estimate the volume of water filtered.

Samples consist usually of a mixture of plankton taxa, fish eggs from various species, algae, jellyfish and other marine fauna and flora. Eggs of the species under scrutiny must

be selected from this mixture. During the ICES mackerel surveys, this might be done either on board or back at the laboratory. If carried out on board, fish eggs are separated visually immediately after the sample is brought on board. The sample is then stored in 4% formaldehyde solution and a second pick takes place 12 hours after the first one. Eggs, or any other material mistaken for them, are also kept in formaldehyde. These are later inspected under a binocular microscope. Eggs from different species are separated according to their morphology, and those belonging to the species being assessed are then staged following standardised descriptions. When egg identification is to be carried out in the lab, the whole sample is stored in formaldehyde after extracting from the net for later analysis.

Egg development in the sea, from spawning to hatching, can be categorised in a series of predictable stages, morphologically recognisable (Gunderson, 1993), although ultimately arbitrary. A variable number of stages are defined for each species, from five in the case of the Atlantic mackerel to ten for South African anchovy (Shelton *et al.*, 1993). The duration of each egg stage is dependent on sea water temperature, and this relationship needs to be established for egg production estimation. For example, eggs of Atlantic mackerel remain in stage I for 57 hours when water temperature is 10°C, and for 22 hours at 18°C. Atlantic mackerel egg production is estimated from counts of stage I eggs only. This stage is divided in two substages, Ia and Ib, and numbers counted separately are later added. Eggs are considered to be in stage I from spawning until the formation of the blastodisc[1], visible as a "signet ring", together with the start of thickening of the egg bundle at one pole (ICES, 2001).

The effectiveness of planktonic surveys at assessing spawning abundance depends in great measure on survey design. Surveys should match the extent, both temporal and spatial, of the spawning event. The grid size applied should reflect the scale at which patterns of abundance in spawning change (Gunderson, 1993). Finally, some level of spatio-temporal replication is required if variance is to be estimated. This is because the current variance estimator (ICES, 1999a) is based on the variability present in samples taken during the same period at the same sampling square.

The usual constraints found on any sampling exercise apply to egg surveys. Limitations in shipping time are usually the most acute of all. Sampling can be severely hampered by bad weather, or interrupted by problems with the equipment. All of these circumstances can severely affect the precision and bias of the final estimate. Bayesian approaches have been applied to the problem of quantifying the bias and precision in the resulting estimate in other studies of fisheries research survey estimates of abundance. The Bayesian approach allows biological, technological and statistical parameters to be treated as random variables. Expert judgement, empirical data analysis and Monte Carlo simulation have been combined under this approach to quantify bias and precision in

[1]Embryo-forming area of egg prior to cleavage

survey biomass estimates. A common procedure involves the estimation of a prior probability distribution for q, the constant of proportionality or scaling parameter. q relates fish or egg abundance, as directly estimated from survey data, to the real abundance once corrected for bias. In an assessment of New Zealand's western stock of hoki *Macruronus novaezelandiae*, McAllister *et al.* (1994) constructed a prior distribution of q, the constant of proportionality for acoustic survey estimates of abundance, from Monte Carlo simulations of the uncertainty in the parameters, such as target strength. A stock assessment of yellowfin sole (*Limanda aspera*) in the Bering Sea applied a similar method to trawl survey estimates of abundance (McAllister & Ianelli, 1997). Prior distributions for q were also constructed by Monte Carlo simulation of uncertainties in survey parameters that affected the effectiveness of the survey in that area. This approach has not been applied previously to abundance estimates from ichthyoplankton surveys.

In this chapter, the possible sources of uncertainty and bias in pelagic eggs surveys are investigated and quantified. A Bayesian approach is then developed to account for the deficiencies in pelagic egg survey design and provide a measure of the reliability and quality of the estimate.

2.2 Data management

Data made available to this project come from the two official repositories established by the International Council for the Exploration of the Sea (ICES) Working Group on Mackerel and Horse Mackerel Egg Surveys (WGMHES). The longest time series comprises the results of triannual surveys carried out since 1977 on the western stock, nowadays considered as the western component of a single western European stock. Its present maintainer is Dr. Dave Reid from the Marine Laboratory Aberdeen (MLA). A second dataset, stored at the Instituto Español de Oceanografía in Madrid, covers the southern component (Spanish and Portuguese coasts). Surveys in this latter area started in 1992, and good coverage was only achieved in 1998. Other data related to the surveys are kept separately, and some are only available through the working group reports.

Data quality is quite variable, having improved notably over the years. Despite this, erroneous entries were found in the dataset, and inconsistencies remain in the storage procedure and database structure. A number of simple measures, following the general guidelines in FAO (1999), could be implemented to improve the management of data from the ICES mackerel surveys, and facilitate their use in research studies, such as:

- Data should be unified into one single dataset, or both datasets should share a common structure and agreed standards.
- Data from all egg surveys should be combined into a single table, with a documented structure.
- Original raw data and calculated data should be distinguished (for example, volume when calculated from flowmeter revolutions and net aperture).
- Data from different years and stock elements are to be stored in the same units.

- A complete database system could be developed incorporating egg survey, fecundity and other survey data and information.
- Agreed procedures could be established for data quality and integrity to be validated. This would include checking for errors in data input.

2.3 The AEP model

In brief, the AEP model quantifies spawning stock biomass from the number of eggs produced by the females in the stock, their average fecundity and the sex ratio of the population. The annual estimate is based on a series of daily egg production estimates for a number of periods over the spawning season, and the annual egg production curve is calculated from those period estimates. The area under the curve is the total annual egg production for the stock.

The first step in the calculation is to estimate daily egg production per unit of surface area for each sampling square from the number of sampled eggs. Daily egg production, D_{psh}, for haul h, in sampling square s, during period p, when based on stage I eggs only, is estimated as

$$D_{psh} = \frac{N_{psh}}{V_{psh}} \cdot \frac{24}{H_{psh}} \cdot d_{psh} \tag{2.1}$$

where N_{psh} is the number of stage I eggs found in haul h of period p in square s, d_{psh} is the sampling depth (m) reached in this haul, V_{psh} is the volume (l) of water filtered during this haul, and H_{phs} is the length in hours of stage I of egg development at the water temperature encountered in this haul. This is in turn calculated according to the equations developed by Lockwood *et al.* (1977) for Mackerel,

$$H = \exp(-1.61 \log(t_{20}) + 7.76) \tag{2.2}$$

and by Pipe & Walker (1987) for Horse mackerel,

$$H = \exp(-1.608 \log(t_{20}) + 7.713) \tag{2.3}$$

where t_{20} is the sea water temperature at 20m depth. An estimate of daily egg production in rectangle s in period p, D_{ps}, is calculated as the mean of all H_{ps} hauls in the rectangle, multiplied by its area, A_s (m^2)

$$D_{ps} = \frac{1}{H_{ps}} \sum_{h=1}^{H_{ps}} D_{psh} \cdot A_s \tag{2.4}$$

If eggs can be assigned to different development stages with sufficient reliability, egg mortality rates can be estimated, and all egg data can then enter the production estimate. If

constant mortality and spawning rates can be assumed during the length of the incubation period, the catch curve method might be applied (Gunderson, 1993). An exponential mortality model of the form

$$N_{hi}/a_{hi} = N_h e^{-Z_{hi}a_{hi}} \tag{2.5}$$

is fitted to the abundance at age (N_{hi}/a_{ij}) and age (a_{hi}, time elapsed from fertilisation to midpoint of stage i) data from haul h and stage i. The intercept, N_h, is the estimated total number of eggs at that location (See Figure 2.1). A constant mortality rate is assumed for all stages. In contrast, the cohort method allows for Z to vary with egg age. Samples need to be taken at the same place during egg development, allowing a single cohort to be followed in time. This method is better suited for boreal species, where egg hatching might occur up to 15 days after spawning. If egg development rates are high, as it is the case for species in warmer waters, the cohort method requires almost continuous sampling over the development period. Extrapolation based on arithmetic means is used

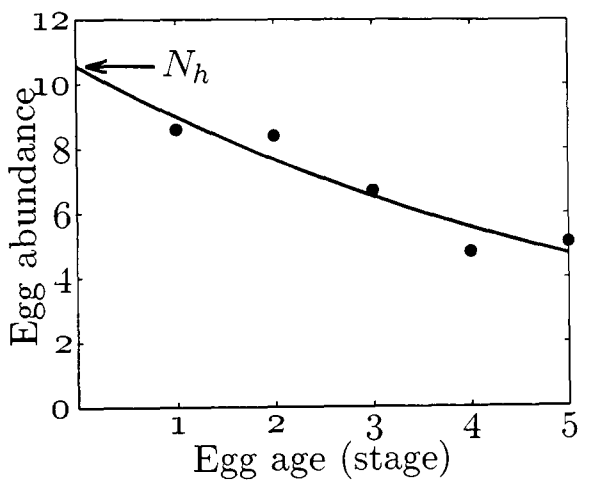


Figure 2.1: Example of catch curve method for estimation of egg mortality (Z) and egg abundance (N_h) from counts of eggs at various stages. See equation 2.5

in unsampled rectangles immediately adjacent to at least two sampled rectangles. Let u index unsampled rectangles, and let $\delta_{nsu} = 1$ if sampled rectangle s in period p is adjacent (either immediately or diagonally) to rectangle u, and be 0 otherwise. Daily egg density in rectangle u is then estimated by

$$D_{pu} = A_u \left(\sum_{s=1}^{S_p} \delta_{psu} \right)^{-1} \sum_{s=1}^{S_p} \delta_{psu} D_{ps} \tag{2.6}$$

where S_p is the number of rectangles sampled in period p.

Daily egg production in each period is then estimated as a sum of the sampled and interpolated rectangles

$$D_p = \sum_{s=1}^{S_p} A_s D_{ps} + \sum_{u=1}^{U_p} A_u D_{pu} \tag{2.7}$$

where U_p is the number of unsampled rectangles adjacent to at least two sampled ones.

Finally, annual egg production, E is estimated by integration of the daily egg production histogram, from sum of the daily egg production estimates in each period p, multiplied by its length in days, λ, and the sum of the DEP estimates in each interpolated period i.

$$E = \frac{D_1}{2}(d_{1,1} - d_0) + \sum_{p=1}^{P}\left[D_p(d_{p,2} - d_{p,1}) + \frac{D_p + D_{p+1}}{2}(d_{p+1,1} - d_{p,2})\right] \quad (2.8)$$

where $d_{p,1}$ and $d_{p,2}$ are the start and end dates (in Julian days) of period p, and d_0 is the start date of the spawning season. For the last sampling period, when $p = P$, D_{p+1} will equal zero, and $d_{p+1,1} = d_P$, the end date of the spawning season. The variance of the annual egg production estimate is the sum of all period variances,

$$Var(E) = \sum_{p=1}^{P} {\lambda_p}^2 Var(D_p) \quad (2.9)$$

which are in turn estimated as,

$$Var(D_p) = CV_p^2 \sum_s \left(A_s + \sum_u \frac{\delta_{us}}{\delta_u} A_u\right)^2 \frac{\bar{D}_s^2}{N_s} \quad (2.10)$$

where CV_p is an estimate of the global coefficient of variation for each period. It is calculated as

$$CV_p^2 = \left(\sum_{p=1}^{P}\sum_{s=1}^{S} H_{sp} - H_s\right)^{-1} \sum_{p=1}^{P}\sum_{s=1}^{S}\sum_{h=1}^{H} (\ln(D_{sh}) - \ln(\bar{D}_s))^2 \quad (2.11)$$

This is based on an estimate of the variance in egg densities for each sampling square (s) with two or more hauls (h) of non-zero values. In most periods the number of squares for which variance can be estimated is very low, so the estimates for all periods are pooled into a single CV.

2.4 Model parameters

A number of parameters are required for the estimation of annual egg production under the AEPM (Table 2.1). An initial exploration was carried out to identify those parameters that are most likely to introduce error or bias in the estimate, and those that can be considered precise and unbiased. Parameters were considered biased if the mean bias was equal or larger than 4%, and as introducing error if the estimate had a CV of more than 5%.

2.4.1 Depth

The maximum depth reached by the net allows the calculation of egg production per surface area of sea from an estimate of production per volume sampled. Depth is recorded

Table 2.1: Parameters present in the AEPM model

Parameter		units
Egg counts	N	eggs
Volume of water filtered	V	m^3
Maximum depth of sample	d	m
Duration ff Stage I	G	Hours
Water temperature	t	C
Area of sampling square	A	m^2
Daily egg production	D	$eggs/m^2/day$

by a probe attached to the body of the net case. Information on depth is transmitted to the boat for monitoring and to decide when the sampler should be raised, and is later stored. One of the most commonly used models, the Scanmar depth sensor, whose precision is reported by the manufacturers to be better than 0.5%.

2.4.2 Stage I length

Development rates of Atlantic mackerel eggs at different temperatures have been determined through a series of tank experiments (Lockwood *et al.*, 1977), where eggs were kept at a range of temperatures, and the time spent at each of the five development stages measured. Results are summarised by equation 2.2. The original data from this experiment were not available, but correlation coefficients (0.993 for stage I) indicate very good agreement with the exponential model in use. Although error was undoubtedly present in the experiment, no data are available for it to be quantified.

2.4.3 Water temperature

As the net is lowered, water temperature is measured at different depths. Temperature sensors are attached to the sampler case. As in the case of the depth sensors, temperature is estimated with high precision, and no significant error seems to be introduced. Bias related to the use of temperature at 20m in the development equation is explored in section 2.5.3 (Page 56).

2.4.4 Area

The area of a sampling square, defined on a 1/2 degree by 1/2 degree grid, varies with latitude. Given the latitude (ϕ_s) of the centre point of square s, area (in m^2) can easily be calculated as

$$A_s = \cos\left(\frac{\phi_s}{180}\cdot\pi\right)\cdot 32^2\cdot 1853.2^2 \tag{2.12}$$

2.4.5 Volume

The volume sampled during each net haul is measured by two flowmeters inside and outside of the net mouth. The inner flowmeter measures the volume of water flowing through the net, while the outside one is used to provide an indication of net clogging or

flowmeter malfunction. To convert the number of revolutions recorded by the flowmeter into a volume measurement, net and flowmeters must be calibrated. This is achieved by recording the revolutions counted when towed over a known distance. Volume, V, in haul h can then be calculated as

$$V_h = \frac{r_h \omega}{c} e \tag{2.13}$$

where r_h is the number of revolutions recorded by the flowmeter, ω is the aperture of the net (m^2), c is the calibration value of the sampler (revs/m), and e is the efficiency factor. This is a measure of the percentage of the water flow directed to the net mouth that is actually engulfed by it. The efficiency is regarded as 100% when the measured volume is equal to the theoretical volume, calculated as the product of the distance towed, and a. e values range from 0.8 to 1.3 for very efficient samplers, although is commonly found to be around 1.

In recent years, some countries have started using computerised systems that provide a single measurement of volume filtered. This trend, likely to spread (D.Reid, *pers. com.*) and the fact that some of the older datasets include only the final volume estimate, forced volume to be investigated and treated as a single variable.

Problems have been identified with the samplers used in the ICES mackerel survey (Brander *et al.*, 1993; Anon., 1999), especially on the efficiency values assumed, and the incorrect calibration of the net-flowmeter-boat system. An European funded project has analysed in detail errors in the calibration and efficiency values of the Gulf III samplers used by most countries in the ICES mackerel surveys. By comparing measurements of flow by the sampler flowmeters with those by an external Laser-Doppler profiler, bias on each of the samplers was quantified. The bias factors uncovered should be applied to previous surveys up to 1998 (Anon., 1999), and it is expected that the new efficiency and calibration factors in use from that year are unbiased. The study recommended that bias corrections should be applied to three of the samplers. Underestimation of the volume filtered by the nets results in overestimation of the egg production, and vice-versa (See equation 2.1). According to the study results, English data up until 1998 should be corrected by -10%, as that is the uncovered overestimation in the volume of water filtered. Data from the Dutch samplers should be corrected by $+16\%$, and the German data should subjected to a -8% correction. Tests were carried out on the effect the introduction of these corrections in the bias model would have on the final estimates. But given the reservations contained in the previous study on the applicability of tank-based measurements to real sampling situations, bias corrections were not included in the final model.

The calibration process must be ideally repeated at the start of a survey, so it reflects the performance of the sampler-boat system. Measurement of the distance travelled and the speed of the boat during calibration tests can be difficult to achieve to a high precision (Anon., 1999), so calibration parameters, might be quite uncertain. No data are available

for quantifying this factor, so an attempt has been made at using expert knowledge to describe the possible error as a probability distribution.

2.4.6 Fecundity

Egg production estimates are by themselves of little use for management purposes, where the quantity of interest is the spawning stock biomass (SSB). Estimation of SSB from egg surveys requires an estimate of female fecundity in the stock. Realised fecundity, in terms of eggs spawned per gram of female body weight, can be divided in two components. First one is potential fecundity, the number of oocytes ready to be spawned present in the ovary at the start of the season. This is assessed by sampling of females prior to the start of the spawning period, and subsequent microscopical counts of the number of hydrated oocytes above a certain size. Due to limitations in food availability and energy needs, a number of oocytes are never spawned, and instead are re-absorbed. This is termed atresia. Realised fecundity, $\bar{F}_r$ is then estimated as

$$\bar{F}_r = \bar{F}_p - \left(\frac{\bar{\alpha}\psi}{l_a} l_s\right) \tag{2.14}$$

where $\bar{F}_p$ is the mean potential fecundity, $\bar{\alpha}$ is the mean relative atresia, ψ is the prevalence of atresia (the percentage of females sampled with atretic eggs), l_a is the length in days of the atresia stage, and l_s is the duration in days of a spawning event.

Estimating fish fecundity is a complex process. First, a sufficient number of mature females must be caught shortly before spawning starts. Ovaries and fish are weighed and measured. These samples must then be stored for a period of 3 months in Gilson's fluid, for oocytes to be released from the ovary tissue. Maturity stage is assessed, and sections are prepared for microscopic examination. If any sign is found that any batch of eggs has already been spawned, the sample must be discarded. Potential fecundity is then estimated by counting the number of oocytes above a certain diameter threshold, previously estimated from histological studies (Emerson *et al.*, 1990). A full review of the methodology presently applied to the ICES mackerel stock can be found in Witthames (2000).

Fecundity estimates can suffer from bias and error arising from different sources. Error is introduced when fecundity slides are analysed, and this can vary greatly depending on the experience of the observer (ICES, 2001). By measuring potential fecundity in terms of number of eggs per unit of body weight, the assumption is made that the relationship between fecundity and body size is a linear one. This might not be the case (P. Witthames, *pers. com.*). Larger fish appear to produce disproportionately larger numbers of eggs, and this is not captured by the mean fecundity estimated from all ages combined. Spatial trends in fecundity and atresia, and temporal trends in atresia are also ignored. Lack of data precludes further examination of these problems.

Uncertainty around fecundity and atresia estimates can be observed from a plot of the distribution of values around the means (Figure 2.2). Variance on the error distribution has decreased over time, from 0.04 in 1992 to 0.02 in 1998. The variation observed might be related to age, size, food availability, energy consumption in the pre-spawning and winter periods, and genetic variability. Some of these factors are likely to change significantly for any season. Consequently, the value should be updated when new information from each survey is available.

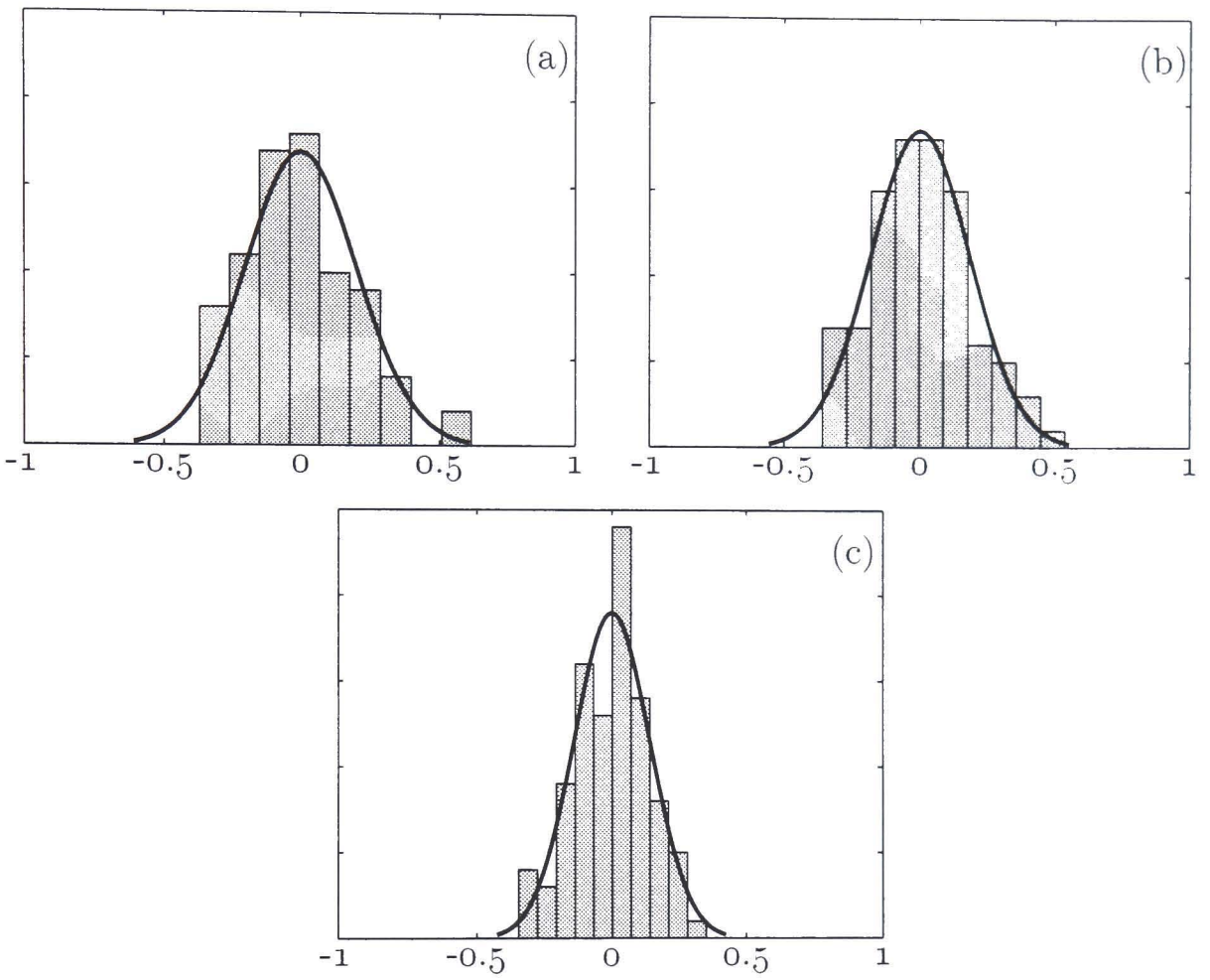


Figure 2.2: Distribution of normalised residuals around mean fecundity estimates for Atlantic mackerel. Data from (a) 1992, n=80, (b) 1995, n=119 and (c) 1998, n=93. Overall distribution is of the form $\epsilon \sim N(0, 0.17^2)$

2.4.7 Egg counts

Together with the error present on any counting procedure, assigning eggs to a development stage poses extra difficulties. The division of the egg development continuum in a finite number of stages is a necessary simplification. Each described stage can be thought of as an average situation that best summarises a range of positions along the development axis. The *distributions* around each stage mean do overlap, that is, the development stage to which an egg should be assigned is not always clear. This difficulty appears to be greater for older eggs. For example, the ICES mackerel assessment is based on stage

I eggs only, as the divergences between observers and laboratories on the status of older eggs are considered to be too large (ICES, 1999a, 2001).

Workshops have been carried out for a number of years to unify criteria and minimise divergences and errors in egg staging. Results from these cross-validation exercises provide a measure of the error introduced at this stage in the process. The same sample of mackerel eggs, initially consisting of 100 of them, is sent around all laboratories involved, and their results compared. The results of the last exchange exercise (ICES, 1999a) are summarised in Figure 2.3. Some disagreement was found between laboratories and observers on the number of eggs to be assigned to stage I in the same sample, sometimes up to a 4% difference. These were even larger at later stages, and also increased for horse mackerel. The most recent workshop conducted (ICES, 2001) found a -2.5% average bias in counts of stage I eggs, with a range from -16% to +15%.

Although the experiment accounted only for error in egg staging, some disagreement was also encountered on the identification of the eggs. In this case, 11 eggs were not identified as mackerel eggs by one of the laboratories (ICES, 1999a, Table 2.1.1). However, egg identification seems to be less problematic than staging and error appears low, as samples are usually processed two or even three times (I. Mosqueira, *pers. obvs.*). Nevetherless, this could be of greater importance for samples with large number of eggs. The distribution of values shown in Figure 2.3 could become wider, although it will still remain a relatively modest source of uncertainty. Bias does not appear to be of great concern. Uncertainty in

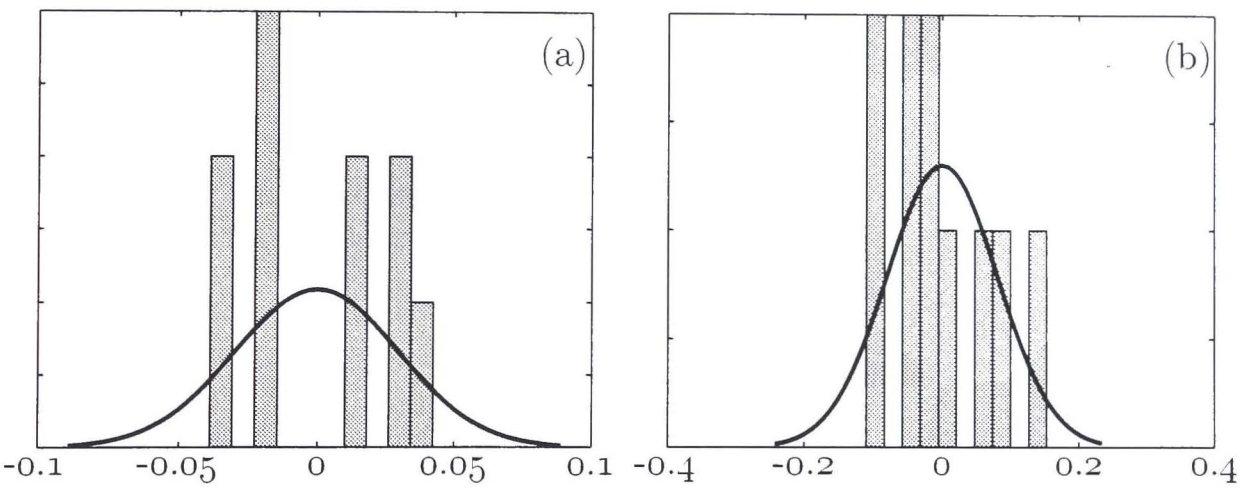


Figure 2.3: Results of the 1998 (a) mackerel and (b) horse mackerel egg exchange exercise. Distribution of the deviations from the mean proportion of stage I eggs in sample ($n = 10$).

the egg staging process, represented by ϵ_s, can then be summarised by normal distributions $\epsilon_s \sim N(0, 0.03^2)$ for mackerel, and $\epsilon_s \sim N(0, 0.08^2)$ for horse mackerel. Although given the limited sample size the normal distribution might appear not to be the best choice for the egg identification error, the process generating these values is undoubtedly accurately expressed by a normal deviation from the real number of eggs in a sample, estimated by the mean value obtained.

Methods are being developed that attempt to provide more precise identification and staging of mackerel eggs and other ichthyoplankton. The use of sample colour, in an informal scale, to assess total egg density has been investigated (Findlay & Lansley, 1998). A method based on the presence and concentration of certain proteins in the eggs is also being tested (P. Morán, *pers. com.*). These methods could provide not only more precise estimates of egg abundance, but also allow the use of all egg stages for egg production and egg mortality estimation.

2.5 Model assumptions and procedures

A number of assumptions, mostly untested, lie embedded in the AEP model, and some of the methods for estimating certain parameters could introduce bias and error in the final estimate (See, for example Hunter & Lo, 1993, for a review of some of the expected limitations of the method). Assumptions were identified and those with potential for significantly deviating the egg production or stock biomass estimates were analysed in detail. These analyses were limited in some cases by the availability of data.

2.5.1 Instantaneous sampling

Egg abundance, as estimated from counts of eggs in the sampling net, are assumed to be an estimate of the number of eggs, of a certain age, at the moment of sampling. Ichthyoplankton nets employ a substantial length of time in sampling the whole depth of the water column, depending on the maximum depth reached. Eggs sampled by the net continue their development during the length of the haul. Consequently, counts of eggs extracted from the net reflect the age of eggs at the moment of fixation, slightly older than when sampled from the sea. A number of eggs might be thus categorised in a later stage than would be appropriate. The significance of this delay is dependent on the ratio between the time needed to complete a haul (G_h) and the length of the development stage (G), which varies with sea temperature.

For Atlantic mackerel in 1998, G ranged from 23 to 73 hours ($\bar{G} = 37.80$), while hauls lasted between 0.25 and 1.4 hours ($\bar{G}_h = 0.56$). The length of the haul represents the maximum time that an egg could spend inside the net, so a mean estimate is given by $G_n = G_h/2$. A Monte Carlo procedure was carried out combining randomly chosen values of G and G_h, to construct a distribution of possible values of the G_n/G ratio. Assuming a constant spawning rate for the length of the haul, the ratio indicates the proportion of sampled stage I eggs in the sample that would have reached stage II by the end of the haul. As Figure 2.4 shows, the bias introduced appears to be very small ($\bar{\epsilon}_i = 0.8\%$).

2.5.2 Net speed and depth coverage

When sampling for planktonic fish eggs like those of Mackerel and Horse mackerel an attempt is made to cover the whole water column, thus correctly estimating egg abundances per unit surface. But limitations in time and resources, together with knowledge of the most likely distribution of eggs with depth, means, in the case of the ICES mackerel

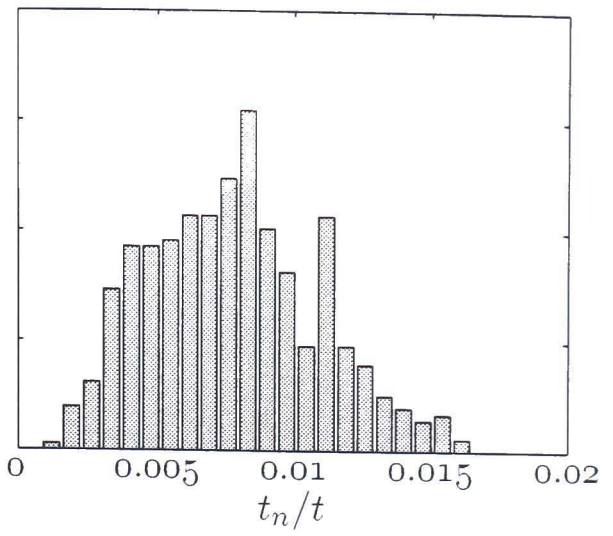


Figure 2.4: Probability distribution of the ratio between length of a sampling haul and length of stage I, t_n/t, for Atlantic mackerel in 1998.

survey, that sampling is limited to the first 200m in deeper waters. At the same time, the deployment process of planktonic nets would ideally be carried out at a constant speed, sampling homogeneously and as close as possible to a *V-shaped* trajectory, at each depth. Departures from this ideal are to be expected, due to the non-static behaviour of the system: boat, net and sea. As Mackerel eggs are not evenly distributed within the water column (Coombs *et al.*, 2001), the combination of both circumstances could lead to significant error and bias in the estimation of egg densities.

The standard procedure for ICES Mackerel surveys is to survey to a close distance from the sea bottom, but limited to 200m meters in waters deeper than that. Effectively, the survey is estimating Mackerel and Horse mackerel egg abundances in a section of the spawning volume (in contrast to spawning area) limited by the latitudinal and longitudinal limits, and a depth limit of 200m. Although most eggs appear to be concentrated above that limit, they are also present in deeper waters in significant numbers, especially at the start of the spawning season (Coombs *et al.*, 2001). Egg abundance estimates might then be underestimating the real value, and this bias could prove more significant if, as the last few surveys have been showing, spawning is moving further west and into deeper waters (Beare & Reid, 2002).

The present analysis attempts to quantify the bias and error introduced in the egg production estimate by the previous questions, and how they correlate with spatial and temporal factors.

Methods

Logs of net deployment were obtained from MLA and CEFAS, in both cases from the 2001 survey. They provided information on the time spent by the net at different depth sections, in the following format: CEFAS, time spent by the net at every ten-meter interval; MLA, depth of the net at every 3s scan. Times from the downwards and upwards sections were

added, and a total of 440 net logs were available. The MLA dataset was transformed to the CEFAS format, and for later comparison with Mackerel egg profiles.

A total of 84 profiles, with egg densities every 5m. and sampled between 1974 and 1991, were obtained from Dr. S. Coombs (Plymouth Marine Laboratory)(Coombs *et al.*, 2001). Taken mostly along the 200m isobath, they stretch from the southern limit of the western component (off the Spanish-French border) to the west of Scotland (off the Isle of Lewis), although they are mostly concentrated between 49 and 51 degrees North, on the shelf of the Celtic Plateau and Bay of Biscay (see Figure 2.5).

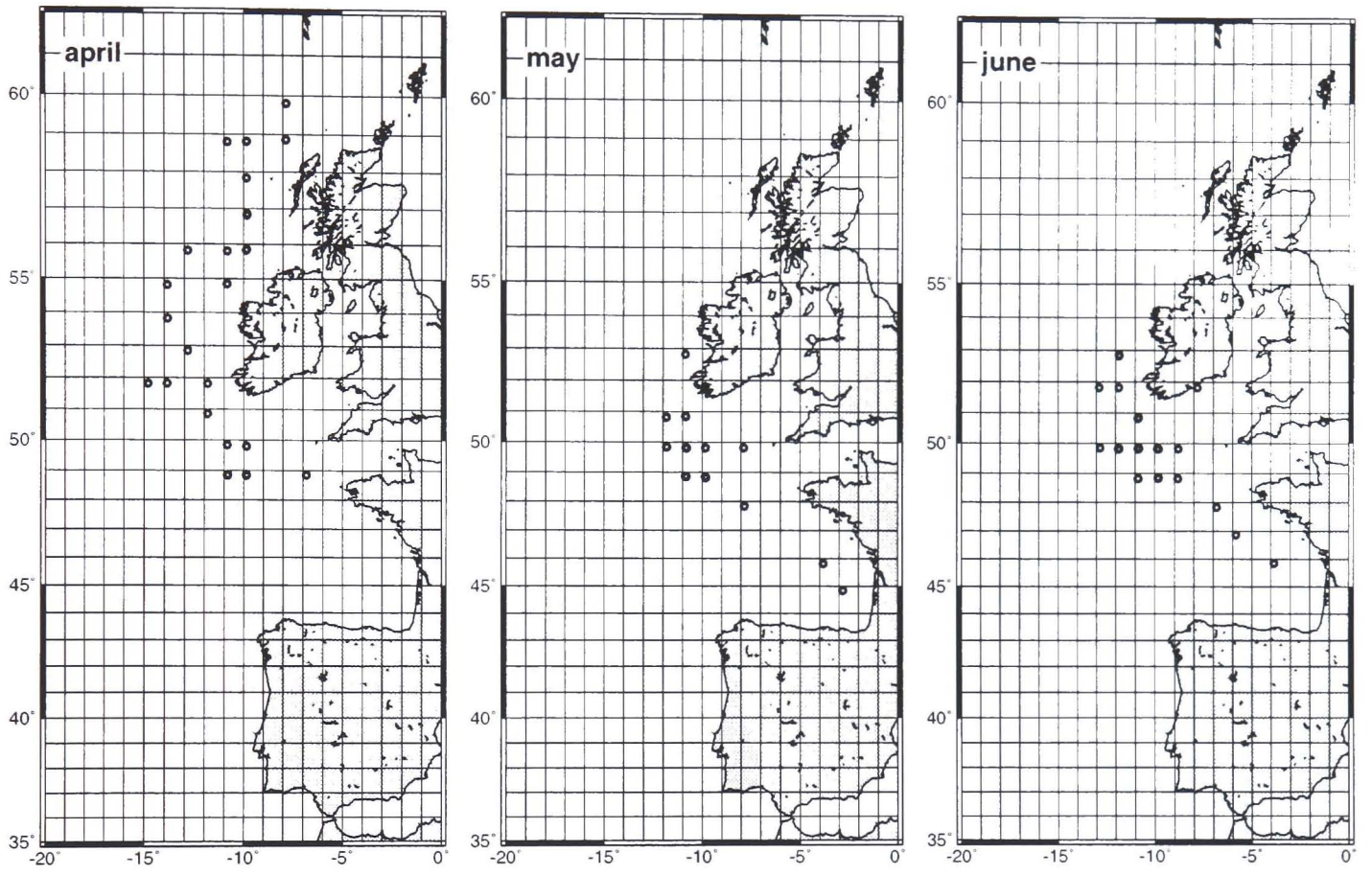


Figure 2.5: Spatial location of egg profiles for the three central months of the survey

The available dataset of Mackerel egg profiles in the survey area shows wide variation in maximum depth of eggs and the pattern of distribution in the water column. Two parameters were used to provide a description of each profile, investigate their variation, and model their relationship with bottom depth (B), latitude (ϕ), Julian date (t) and egg numbers (N). The expected egg depth ($\bar{d}$) is defined as the depth where egg density is highest, that is the mode of the egg distribution with depth. The second is the percentage of eggs found below 250m (N_{250}) in those profiles that reached deeper than that threshold. A higher value than the agreed limit (200m) was used to allow for a more conservative estimate, and because samples taken in waters between 200 and 250m are usually extended to the bottom, 250m appeared to be a more realistic, and conservative, limit. The relationships between those variables was analysed by means of linear models (McCullagh & Nelder, 1989), with a normal distribution assumed for all response variates.

To estimate the error and bias introduced by the behaviour of the net, egg densities calculated following two different methods were compared. First, egg densities at each depth section, N_k, were multiplied by the time the net could be expected to stay at each depth section if sampling was perfectly equal, T_k, where $T_k = \frac{T_h}{K}$. T_h is the total sampling time in haul h, and K the number of sections in that haul. The expected, or real, total egg densities were thus estimated as

$$N_h = \sum_{k=1}^{K} \frac{N_k T_k}{T_h} \tag{2.15}$$

This was compared with an estimate where the actual times spent by the net at each section were used instead. The actual total egg densities, $\hat{N}$, were similarly obtained as a mean of the section egg densities, weighted by the time spent by the net on each section, T_k

$$\hat{N}_h = \sum_{k=1}^{K} \frac{N_k \hat{T}_k}{T_h} \tag{2.16}$$

The error between N and $\hat{N}$ was then calculated as

$$\epsilon = \frac{\hat{N} - N}{N} \tag{2.17}$$

whose mean or expected value , $\bar{\epsilon}$, is the bias factor. The root mean squared relative error, RMSRE, was calculated as

$$\text{RMSRE} = \sqrt{\frac{1}{n} \sum_{k=1}^{K} \frac{\left(\hat{N}_k - N_k\right)^2}{N_k^2}} \tag{2.18}$$

where n is the number of Monte Carlo replicates. Finally, the coefficient of variation, CV, was computed as

$$CV = \sqrt{\text{RMSRE} - \bar{\epsilon}_n^2} \tag{2.19}$$

Two different sources of bias were identified, and two similar analyses were conducted to investigate them. First, error arising due to the non-perfect behaviour of the net (ϵ_n) and second, error arising because of the non-sampling of eggs present at depths greater than 200m (ϵ_d). To estimate the first one, egg profiles were matched with time profiles of similar depth according to the following procedure: first, profiles were divided according to their depth, from 10 to 1000 meters, in 50 meters steps, with all egg profiles over 250m deep falling into the same group. Second, a time profile was randomly chosen for each egg profile from the equivalent depth section. As the intention of this analysis was to quantify the error introduced by the imperfect behaviour of the sampling net, only the first 200m meters were used in all deeper egg profiles. Equation 2.17 was then applied.

To estimate the second source of error, ϵ_d, the percentage of eggs below 200m in those samples taken in waters deeper than $200m$, was computed.

$$\epsilon_d = \frac{N_s - N}{N} \tag{2.20}$$

where N and N_s are the total number of eggs in the water column and the number in the sampled 200m upper layer, respectively. This, together with the number of samples on both areas ('deep', $\leq 200m$, and 'shallow, $> 200m$), the surface of both areas, and the contribution of each of them to the total egg production, allow for an assessment of the error introduced in the final AEP estimate. The relationship of both errors, ϵ_n and ϵ_d, with latitude, depth, Julian date and number of eggs in the sample was investigated by multiple regression and analysis of variance.

Results

The data in the egg profile dataset, although not totally representative of all possible situations in the survey area, captures a wide range of them and allows an estimate to be made of the bias introduced in the AEPM procedure. Geographical range is fairly wide, covering from 45° N to 55° N, most of the spawning area of the Western component of the stock. The range of depths covered goes from 33 to 1000m deep. Samples were taken over a number of years (from 1974 to 1987), ensuring greater variability in weather and environmental conditions.

The available dataset of Mackerel egg profiles in the ICES Northeast Atlantic area is fairly limited in number of samples and spatial distribution (See Figure 2.5).Of a total of 84 egg profiles used in the analysis, 41% of them were located in waters deeper than 250m, and 21% contain eggs in layers deeper than 200m. That is, 53% of deep water egg profiles found eggs at depths usually left unsampled when surveying for AEPM estimation. The percentage of eggs in those profiles present below 250m varies greatly ($\bar{N}_{200} = 27.6$, $SD = 25.9$, $n = 18$).

Plots of all variable combinations in Figure 2.6 show some interesting patterns. First, a relationship between $\bar{d}$ and the three sample variables appears to be present, positive in the case of d and ϕ, and negative for J and N. In contrast, eggs appear to be found in deeper waters at the start of the season only (Figure 2.6f), at high latitudes (2.6d), and in an area around 400m bottom depth (2.6b). There is a significant linear relationship between $\bar{d}$ and both bottom depth and date (correlation coefficients of 0.78 and 0.82 respectively). The centre of gravity of the egg distribution tends to move towards deeper waters as the water column expands, albeit at a slower pace. In contrast, eggs appear to be more abundant at shallow depths as the season advances. Surprisingly, the proportion of eggs below 250m does not increase with bottom depth, which in these profiles is equal to the sampling depth. Bottom depth of profiles with highest percentages of eggs below the 250m threshold is mostly concentrated around the 250 to 500m range.

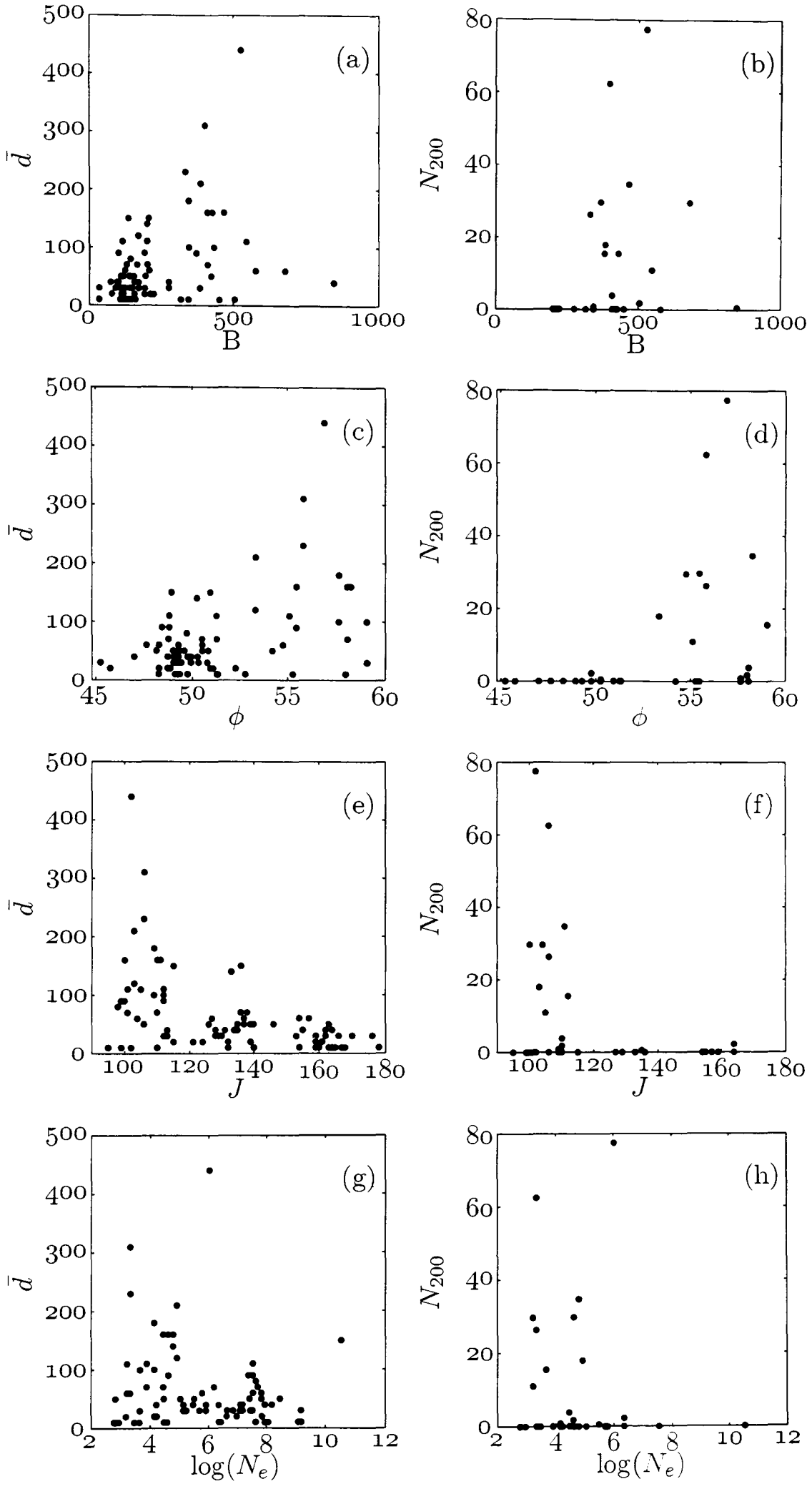


Figure 2.6: Plots of the relationship between depth in meters where most eggs are concentrated, ($\bar{d}$), and the percentange of eggs found below 200m (N_{200}), against latitude (o), bottom depth (B), Julian date (J), and egg numbers in the sample N_e in log scale.

The possible bias introduced by the variations in vertical speed of the net (ϵ_n) does not appear to be of great significance (RMSRE = 0.01). The time spent by the samplers at different depths does not seem to follow any regular pattern, with the exception of a small tendency to undersample close to the surface. The error generated by variations in the speed of the net is introduced in the estimate of egg abundance, in common with any other sampling procedure.

In contrast, egg abundances estimates appear to be significantly biased by limiting sampling to the upper 200m section of the water column (RMSRE = 38.3%). In a small number of profiles, up to 80% of the eggs were concentrated below 200m (Figure 2.7). The relationship between ϵ_d and the four independent variables (J, ϕ, B, and N) was investigated by multiple linear regression. The *full* model showed no significant correlation between the dependent and the independent variables, but some of the individual variables did, as expected after an inspection of Figure 2.6. Both date and latitude provided very good fits to the data when analysed separately (J: $\beta = -0.048$, $t = -3.904$, $p < 0.001$; ϕ: $\beta = 0.264$, $t = 4.404$, $p < 0.001$), but not when inserted in the model together. The correlation between both variables is nearly significant, and appears to be strong enough to influence the model. Changes in the proportion of eggs below 200m with date are expected, due to changes in oceanographic conditions, wind mixing and stratification. Although a linear model provides a significant fit, a look at a plot of date versus ϵ_d shows how profiles with large errors concentrate at the start of the season, and how variance in ϵ_d decreases with date. Error values before and after the 1st of may ($J = 120$), were significantly different (ANOVA, $F_{1,29} = 13.21$, $p = 0.001$). Error in those samples taken before the 120th Julian day conformed to a quite singular probability distribution (Figure 2.7). An uniform distribution between the maximum and minimum values ignoring zer ignoring zero was chosen to represent the variability in ϵ_d for samples with $J \leq 120$. Error was considered negligible for samples after that date.

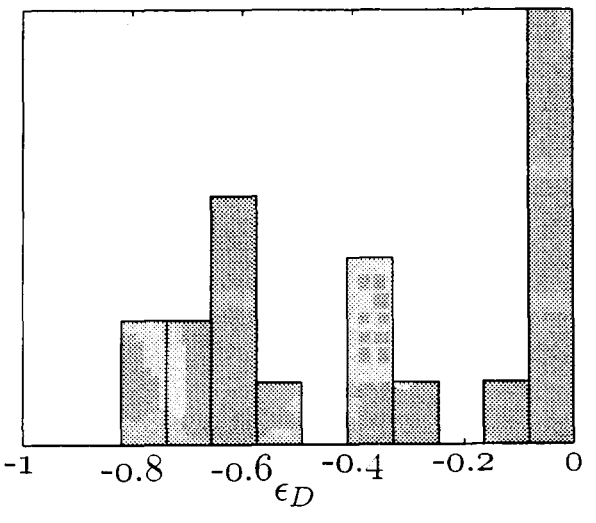


Figure 2.7: Distribution of vaues for the error due to the 200m sampling limit, ϵ_d, for samples taken before the 1st of May (n=31).

2.5.3 Use of t_{20} as egg development temperature

Estimates of daily egg production in the ICES mackerel AEPM are based on counts of the numbers of mackerel eggs in the first stage of its development, or stage I eggs (Lockwood, 1988). The length of stage I allows for the conversion of egg counts into daily egg production, and is dependent on the temperature the eggs experience in the sea. The relationship between water temperature and stage length was established by Lockwood *et al.* (1977) by observation of eggs incubated in an aquarium tank. Similar work was conducted for Horse mackerel (Pipe & Walker, 1987). When these equations are applied to egg counts from the AEPM, the temperature at 20m depth (t_{20}), as recorded by the sampling net, is used as the best representation of the conditions in the water column. The assumption is that t_{20} is an unbiased representation of the conditions encountered by the eggs anywhere in the water column. Mackerel eggs are almost completely neutrally buoyant (Coombs *et al.*, 2001), so their development, at least during stage I, is likely to take place at spawning depth. If a significant number of eggs were to develop at temperatures sufficiently different from t_{20}, bias could be introduced in the daily egg production estimate.

In order to quantify the possible bias involved, an analysis was carried out by using full temperature and egg profiles, whose daily egg production results were then compared with those achieved following the traditional method. The correlation between error introduced by the use of t_{20} and depth, latitude, date, and number of eggs was investigated through linear modelling.

Methods

Data were compiled on mackerel egg distribution with depth and temperature profiles from the same area. The latter were extracted from the Global Ocean Database[2], using the Ocean Data View (ODV 4.1) software. Data were chosen from the survey area only (between 40° and 60° N, and 1° and 15° W), and precise location, date and bottom depth were recorded. A total of 2043 temperature profiles were available, covering depths from 50m to over 2000m. They were recorded over a number of years (1960-1998) and should, therefore, include a wide range of oceanographic and meteorological conditions. Mackerel egg data were again obtained from Dr. Steve Coombs (Coombs *et al.*, 2001). Description of the dataset can be found in page 50.

The analysis matched egg numbers and water temperatures at 10m strata and then computed the total egg production. At the same time, a traditional estimate was carried out by using the temperature at 20m from the same temperature profile. Only temperature profiles on areas at depths smaller than 500m were used. This is the maximum depth reached by the egg profiles. The temperature profiles are structured in sections of unequal sizes, so the egg data had to be transformed from their original 5m-deep strata. The resultant structure is summarised in Table 2.2.

[2] http://www.nodc.noaa.gov/

Table 2.2: Depths covered on each section of the egg and temperature profiles. Upper limits on the egg profiles were part of the lower section.

Temperature	*Egg abundance*
0	$0-5$
10	5-15
20	15-25
30	25-35
50	35-60
75	60-85
100	85-110
125	110-135
150	135-160
200	160-225
250	225-275
300	275-350
400	350-450
500	450-500

Egg profiles were matched with randomly selected temperature profiles from the same depth range and month. The difference in egg production as calculated with both methods, using t_{20} ($D_{t_{20}}$) and the whole profile (D_p), was computed as

$$\epsilon_t = \frac{D_{t_{20}} - D_p}{D_p} \tag{2.21}$$

from which RMSRE and mean bias was computed. Error ϵ_t is calculated by randomly sampling from the temperature dataset, so different runs would yield slightly different results. The procedure was therefore repeated for 1000 times, to generate a long-run distribution of error values. Linear models were fitted to analyse the variation of ϵ_t with depth, latitude, date and number of eggs. Error values, ϵ_t, were transformed to $\log(\epsilon_t + 1)$ to comply with normality and homoscedasticity assumptions. Models were fitted to the values of $\log(\epsilon_t + 1)$ from 500 iterations of the error calculation, and the significance of the regression coefficients was estimated as the proportion of model runs on which the variable was found to be significant.

Results

Bias introduced by the use of the temperature at 20m, t_{20}, as representative of the whole water column for egg production estimation, appears to be important (RMSRE = 14.85%, CI = 13.22 – 22.87). A model of the form $\log(\epsilon_t + 1) \sim J + \phi + B + N$ was fitted to the data. After transformation, $\log(\epsilon_t + 1)$ was normally distributed, and no significant autocorrelation was found, as indicated by a Durbin-Watson test. None of the cross-correlation coefficients among the independent variables were significant (Figure 2.8).

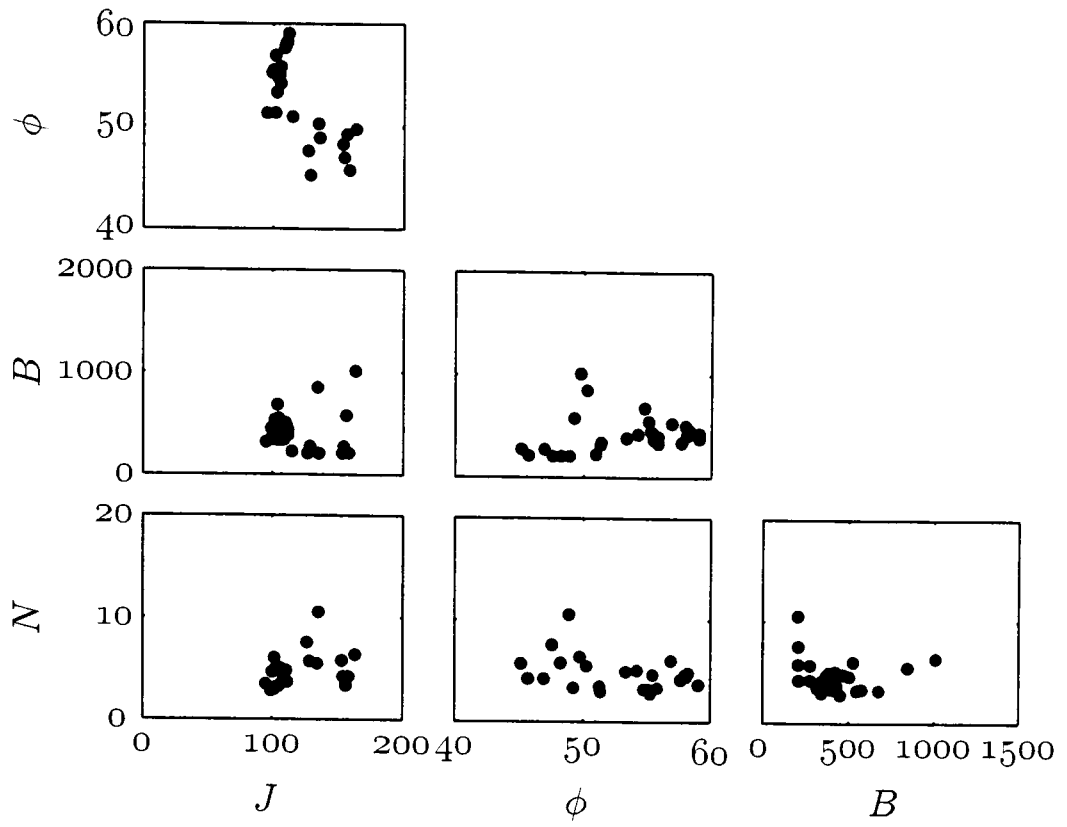


Figure 2.8: Correlation plots between the explanatory variables in the t_{20} and depth coverage analyses: latitude (ϕ), bottom depth (B), number of eggs (N), and Julian date (J). Pearson correlation cofficients were all non-significant.

A *full* model, with the four independent variables included, showed date to be the only variable consistently significant ($p = 0.023$). A simplified model was therefore tested ($\log(\epsilon_t + 1) \sim J + \phi$). Date appears again to be the only variable linearly related to t_{20} error. A simple linear regression modelled the relationship between error and date as $\log(\epsilon_t + 1) = -0.091 - 0.0018J$. Both coefficients were highly significant ($p < 0.0001$). ANOVA results also indicated a good fit to the data ($\bar{F}_{1,82} = 30.15$, $p < 0.0001$), consistent in all 500 replicates. Results were then stratified according to depth, in two strata defined around the 200m isobath. This is the depth at which sampling usually stops, so samples taken in waters shallower than 200m should not show this error, as sampling continues to the bottom. Linear models of the form $\log(\epsilon_t + 1) \sim J$ were fitted to both subsets. Error in samples with depths down to the 200m isobath were modelled as $\log(\epsilon_t + 1) \sim 0.291 - 0.0017J + N(0, 0.068^2)$, while error in those in waters deeper than 200m followed the model $\log(\epsilon_t + 1) \sim 0.417 - 0.0023J + N(0, 0.12^2)$, where the normal distributions model the regression residuals.

2.6 Bias model

Bias has been identified in two of the AEPM procedures: the use of temperature at 20m as egg development temperature (ϵ_t), and the limit of sampling to the upper 200m layer (ϵ_d). A bias model was created incorporating both sources, and their dependency on Julian date and bottom depth. Samples on deep waters (over 200m bottom depth) and taken in the first half of the spawning season, were corrected for error due to the use of t_{20} and for the

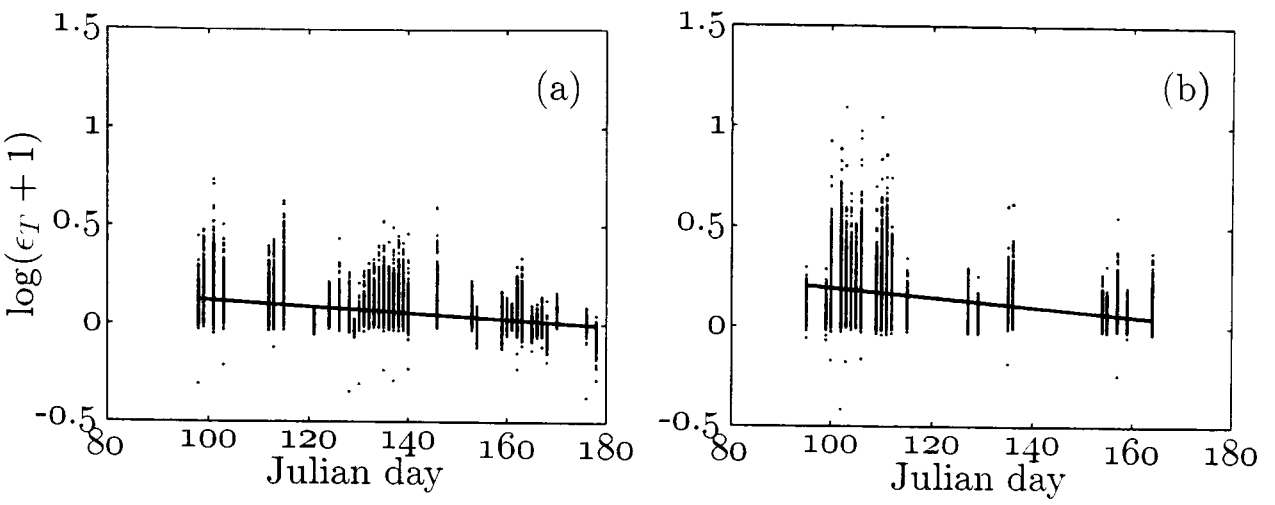


Figure 2.9: Model of change in bias due to the use of t_{20} with time, and estimated errors from 500 Monte Carlo iterations, for samples in waters shallower (a) and deeper (b) than 200m. Fitted lines are $\log(\epsilon_t + 1) = 0.291 - 0.0017J$ and $\log(\epsilon_t + 1) = 0.417 - 0.0023J$, respectively.

proportion of eggs left unsampled below the 200m layer. The first source of bias, t_{20}, was estimated for each sample from a two-step linear model on Julian date and depth, with associated lognormal error. The second one was estimated from an uniform distribution between 0 and -0.8. Only the t_{20}-related correction applied for all other samples not included in this group. The complete procedure, as summarised in Figure 2.10, consisted of the following steps

1. The traditional DEP estimate (D) is produced from survey data for each sample.
2. Samples are selected depending on bottom depth, and error due to t_{20} is calculated. Different models are applied to samples on either side of the 200m isobath. Both models are of the general form $\epsilon_t \sim \exp(\beta_0 + \beta_1 J + N(0, \sigma^2)) - 1$.
3. DEP estimates, corrected for t_{20} bias, are calculated as $D'_c = D/(\epsilon_t + 1)$.
4. Samples in waters deeper than 200m and before the 120th Julian day are selected for correction due to 200m sampling limit.
5. Bias values are randomly sampled from an uniform distribution, $\epsilon_d \sim U(-0.8.0)$.
6. Samples affected are corrected again, using the t_{20}-corrected DEP estimates, $D_c = D'_c/(\epsilon_d + 1)$. For all other samples, ϵ_d equals 0.
7. The ratio between D_c and D for each individual DEP estimate is calculated.
8. The process is repeated 500 times to generate a distribution of values for D_c.

The bias model was applied to the 1998 ICES survey data, including data from both the western and southern components of the stock. Bias on previous surveys will be analysed in Chapter 3, together with the estimates of the influence of the bias model on the final AEP estimates.

Results

The ICES mackerel AEPM appears to overestimate daily egg production in 1998 by an average of 3% (i.e. $D_c/D = 0.97$), although the distribution of values shows that bias of

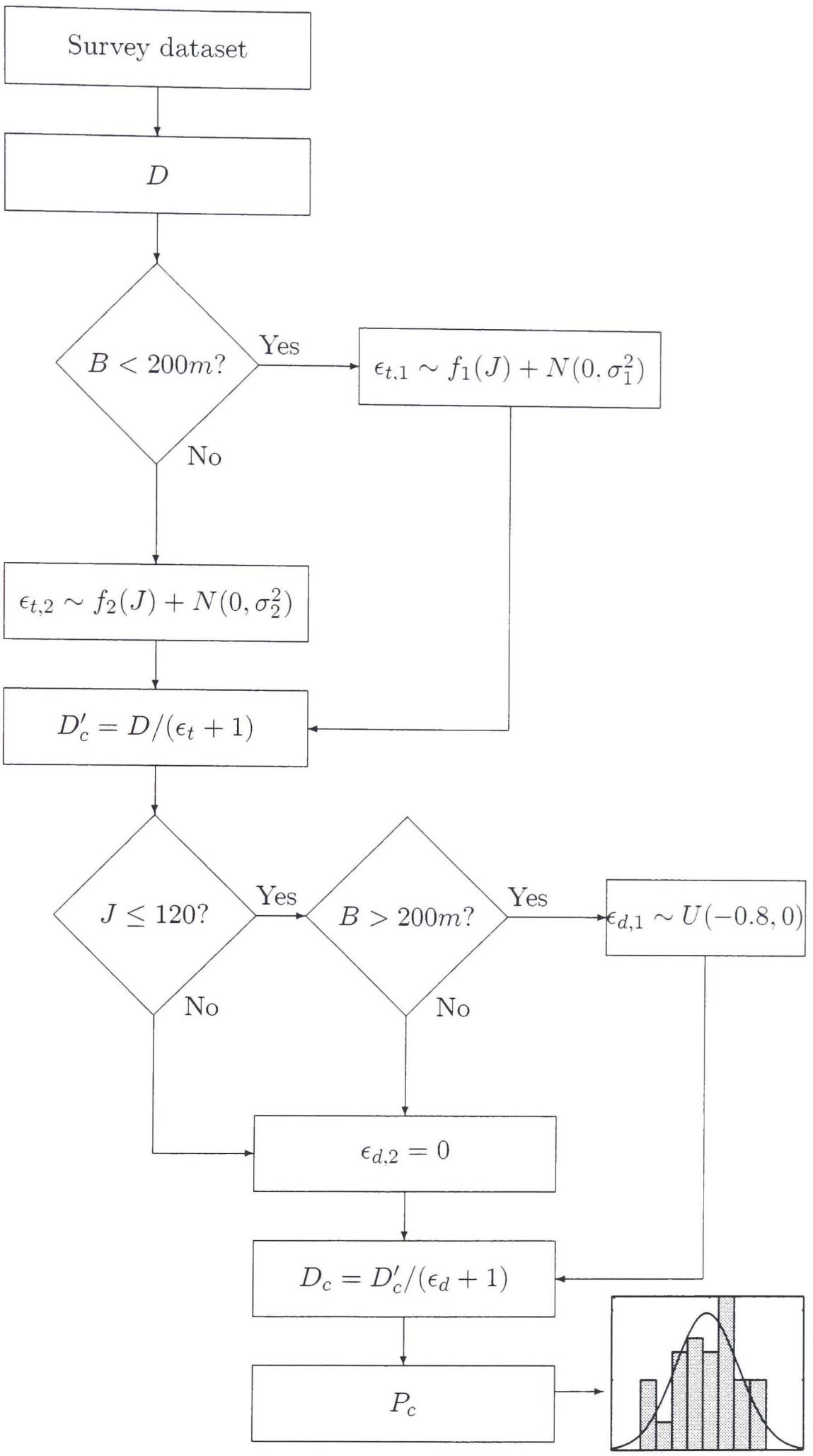


Figure 2.10: Procedure for the application of the bias model and estimation of q, the constant of proportionality. ϵ_t, the bias due to the use of t_{20}, is a function of Julian date, of the form $f(t) = \exp(\beta_0 + \beta_1 t) - 1$, with different parameter values for samples on waters deeper or shallower than 200m. All other variables are defined in the text, except D'_c, the DEP corrected for the first source of bias, D_c, the bias-corrected DEP, and A_c, the bias-corrected AEP.

individual DEP estimates could range from -25% up to +450% (Figure 2.11). Extremely high values of the D_c/D ratio are caused by the small percentage of samples with large proportions of eggs below the 200m sampling limit. The jagged shape of the central part of the distribution is caused by the temporal irregularity of the samples, concentrated on certain sampling periods, thus covering a limited range of date values. Despite the relatively low average value, bias in annual egg production could be more important if samples with a large number of eggs were to fall on the extremes of the distribution. For example, in 1998, the 2% of samples with higher egg production amounted to over 50% of the sum of daily egg production. Large bias in a reduced number of samples would therefore greatly influence the final SSB estimate, and increase the uncertainty in its precision.

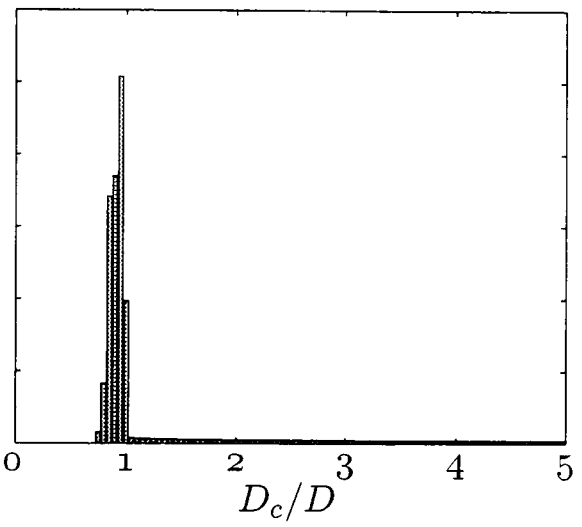


Figure 2.11: Distribution of the ratio of bias corrected (D_c) to traditional DEP (D) estimates, for the 1998 ICES mackerel survey data and on each individual DEP estimate. Result of 500 Monte Carlo simulations. Probability is non-zero for ratio values up to 4.87.

Each of the two sources of bias contribute differently to the final bias ratio. On average, DEP estimates incur a -4.6% bias because of the 200m sampling limit. Despite the potentially large values of ϵ_d, only a small number of samples are affected, 13% of them in 1998, so its overall impact on the sum of daily egg productions is much reduced. But higher percentages of egg production were found in deep waters in previous years (Table 2.3), so larger biases might have been introduced in the AEP estimates. After the correction for t_{20} error is applied, DEP estimates are, on average, +3% biased. This second source of bias originates as eggs are assumed to develop at higher temperatures than is actually the case. Although the range of values of ϵ_t is narrower than for the previous factor, it applies to all samples, and subsequently influences more significantly the final DEP estimates.

The four strata, defined as before and after 120th Julian day. and on both sides of the 200m isobath, show moderately different results (Figure 2.12). Larger biases are present

Table 2.3: Percentages of the daily egg production and sampling sites found on both sides of the 200m isobath.

		1986	1989	1992	1995	1998
% Egg production	$< 200m$	11.0	27.9	36.8	26.0	48.5
	$> 200m$	91.7	69.9	61.2	72.8	49.2
% Samples	$< 200m$	58.6	53.2	56.0	43.9	43.7
	$> 200m$	40.0	45.1	43.0	54.9	55.5

in those samples taken in the first half of the spawning season (top row). Egg productions in either depth stratum are significantly different ($F_{1,626} = 379.5$. $p < 0.0001$). but not between both parts of the spawning season. The pattern in the distribution of values in Figure 2.12a shows the effect of the existence of gaps in coverage between periods. Also. a small number of samples taken at the very start of the season in the Southern area are assigned very large error values by the t_{20} model ($E_c/E \leq 0.8$). The bias model might be overestimating for those samples, as they fall substantially outside of the range of dates used to generate the model fit. But only eleven samples are affected by this problem, all of them with low egg production values, so their overall influence should be minimal. Error arising because of the 200m sampling limit only affects samples taken on the first half of the season and in depths greater than 200m (Figure 2.12b). The combined influence of both sources of bias causes these values to have significant bias (mean $E_c/E = 1.68$), although their overall influence in the 1998 estimate of AEP is small. In 1998, that category accounted for 3.5% of the total daily egg production and 5.4% of all samples. If higher percentages of the spawning were to take place in deep waters and on the first half of the season, AEP estimates could be very significantly biased. For example, the AEP curve for the ICES mackerel survey in 1992 appears to be bimodal, with large numbers of eggs sampled during the first half of the season, and 37% of the egg production occurring in deep waters. This combination of factors could lead to a large bias in the AEP and stock biomass estimate.

The distribution of values of the corrected DEP for samples in deep waters and for the first half of the season, mimics that of the sampling limit corrected values only (Figure 2.13a), as ϵ_d clearly outweighs error from t_{20}. Values corrected for the ϵ_t only, follow the pattern found in the other three strata, the two modes due to the two periods of sampling before the 120th Julian day.

Similar differences can be found among the distributions of ϵ_t for the four strata (Figure 2.14). Again, the jagged distributions due to the patchy temporal distribution of values in the survey are present. The distribution for the stratum with a higher number of samples, which includes samples taken early in the season and in shallow waters (Figure 2.14c), shows clearly how as sample size increases, results tend to follow more closely a normal distribution. As expected, the values of ϵ_d drawn from the $U(-0.8.0)$ distribution mirror

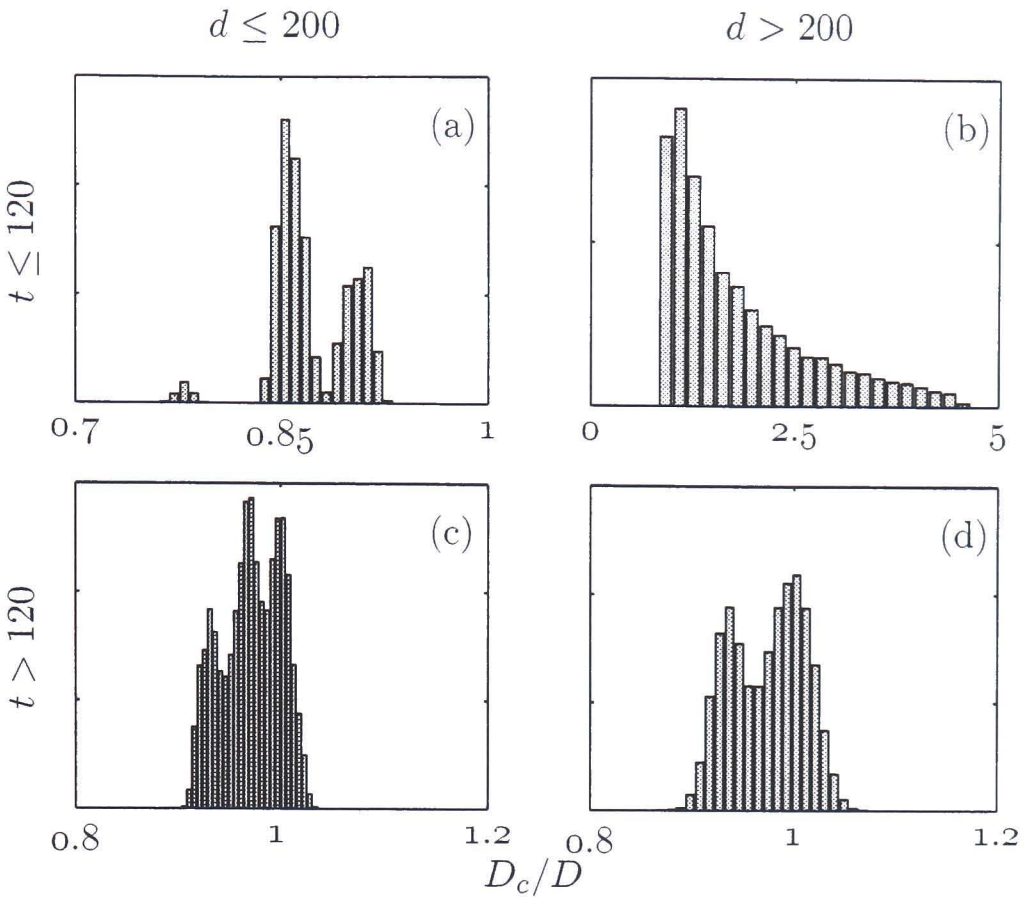


Figure 2.12: Distribution of the ratio between bias corrected, D_c, and traditional, D, daily egg production estimates, for the four time and depth strata. Result of 500 Monte Carlo simulations on the 1998 survey data. Note differences in axis limits.

the pattern found in the experimental data (Figure 2.7), as no environmental variable has any influence on it.

The test results shown above indicate the importance of two basic assumptions present in the ICES mackerel AEP estimation procedure. Bias in the 1998 survey appears to be significant (Figure 2.15) for a number of samples, although the overall effect in the final biomass estimate should not be large. Similar tests should be made on the datasets from previous years, as changes in the spatial and temporal patterns of spawning could change substantially how these sources of bias affect the AEP estimate.

The bias model will also allow the estimation of a constant of proportionality or scale parameter, q, for all ICES mackerel surveys. This will be carried out by comparing a traditional AEP estimate (P) with a bias-corrected distribution of estimates (P_c). The latter estimates are constructed by applying any of the AEPM estimators (traditional, GAM or geostatistics) to the 500 replicates of the survey data, corrected by applying the bias model.

2.7 Conclusions

The analyses presented have investigated and quantified the possible sources of error and bias in standard implementations of the annual egg production method. The bias model constructed shows how the influence of model assumptions can be quantified and applied to ichthyoplankton survey data. This study has been based on one of the earliest examples

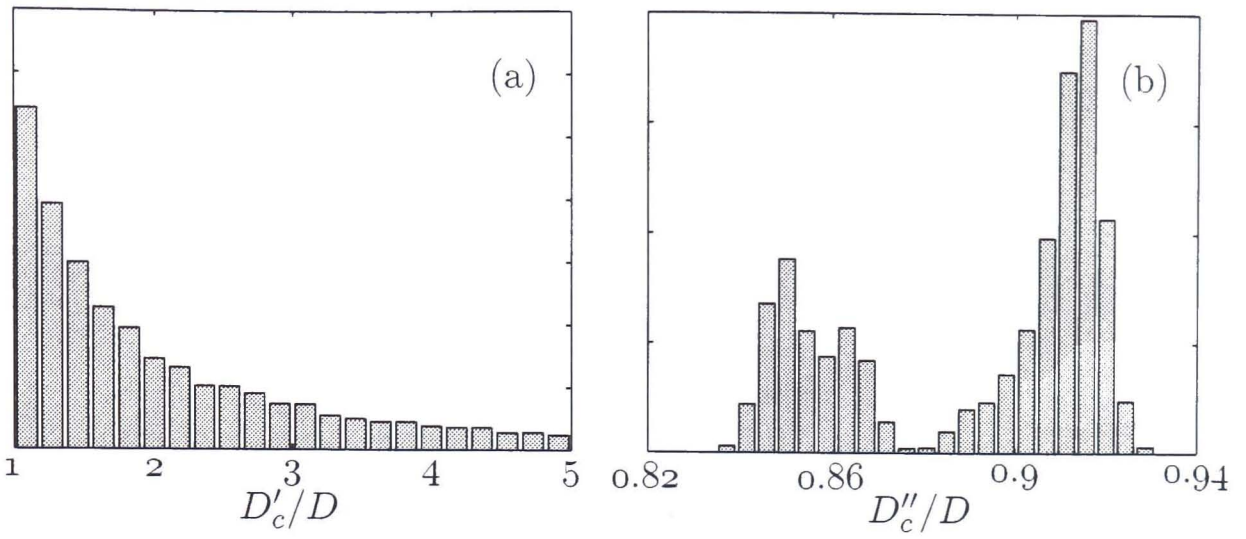


Figure 2.13: Distributions of the ratios between depth bias-corrected, D'_c, and t_{20} bias-corrected, D''_c, against the traditional, D, daily egg production estimates. for samples before the 120th Julian day and in waters deeper than 200m. Combination of both corrections yields D_c, the final corrected DEP estimates. Result of 500 Monte Carlo simulations on the 1998 survey data.

of the use of the AEPM, the egg surveys carried out by ICES, for estimation of mackerel and horse mackerel biomass on the European Atlantic coast, from 1977.

Two of the sampling procedures introduce bias in the estimate. First, by limiting sampling to the range of depths assumed to contain most of the spawn, significant numbers might be left unsampled, if and when oceanographic conditions and adult behaviour mean than waters below the sampling limit contain non-negligible concentrations of eggs. For northeast Atlantic mackerel, egg production is being underestimated as potentially large numbers of eggs are left unsampled in the early part of the season. Bias is introduced in daily egg production estimates in the first half of the season, when, due to the prevalent oceanographic conditions, eggs can be found in depths down to 500m. The same problem has been encountered in other applications of the AEP method. Surveys of walleye pollock off the Alaskan coast are carried out down to 200m deep only, but eggs are negatively buoyant and known to occur in deeper layers (Kim & Gunderson, 1989). Final egg production estimates were thus corrected for the proportion left unsampled, estimated from sequential sampling of the water column. In areas of modest maximum depth the whole water column is sampled (Bulman *et al.*, 1999; Armstrong *et al.*, 2001), but this might not be a viable option in all situations.

Second, the use of sea water temperature at 20m depth in the estimation of daily egg production (Equation 2.2), biases the egg production estimate as eggs are assumed to develop at higher rates than it is actually true. Egg production per time unit is estimated according to egg development equations that typify how the time it takes to move to the next stage and finally hatch is related to water temperature. Th convention in ICES mackerel surveys is to use the temperature at 20m, t_{20}, as representative of the conditions encountered by all eggs present in the water column. The present analysis suggests this might not be a reasonable assumption. Conditions along the section of the water column

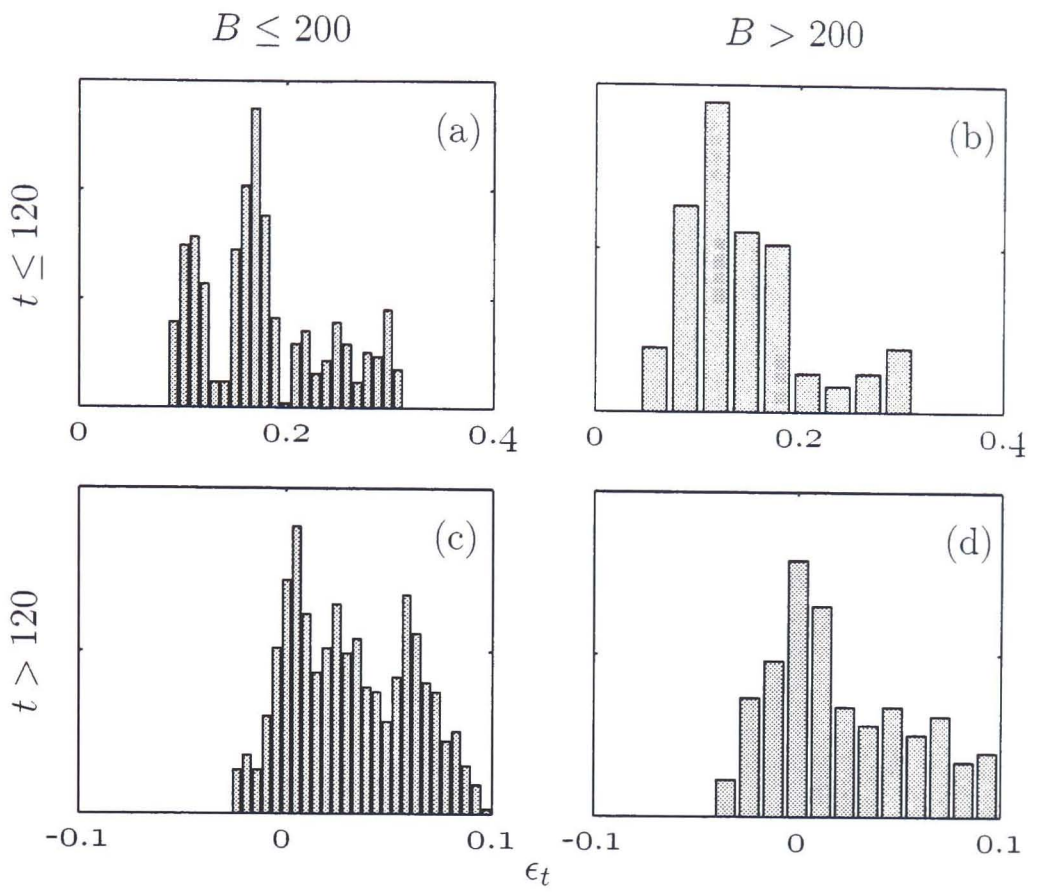


Figure 2.14: Distribution of bias arising from the t_{20} procedure, ϵ_t, from 500 iterations on the 1998 ICES survey data, and on the four depth and time strata.

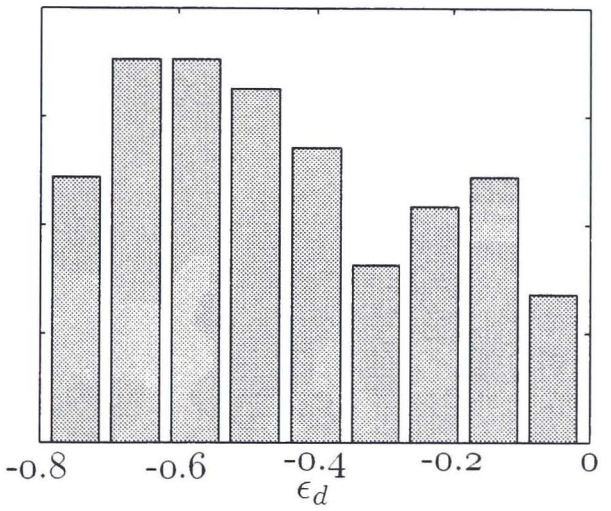


Figure 2.15: Distribution of bias from the 200m sampling limit from 500 iterations on the 1998 ICES survey data. Values of ϵ_d were drawn from a $U(-0.8, 0)$ distribution.

occupied by eggs do vary significantly from those at 20m depth. Temperatures towards the bottom are substantially lower, and egg numbers at those depths are large enough for a significant bias to be present. Bias is lower towards the end of the spawning season, as the oceanographic conditions, lack of oxygen below the thermocline, force adult mackerel to spawn in the shallower layers. When stratification appears, eggs are concentrated in a narrow band of about 100m from the surface, although cases have been observed of extreme restriction of eggs to the 10 or 20m upper layer (Coombs *et al.*, 2001). Limited information is available on the distribution of mackerel eggs in the absence of stratification.

but the indications are that they are present in significant numbers at depths down to 500m.

The problem of precise egg development temperature has also being identified in other studies. Bulman *et al.* (1999) used two different temperature scenarios when assessing AEP of blue grenadier off western Tasmania. Although blue grenadier eggs are positively buoyant, no reasonable assumption could be made on their depth distribution, so the extreme values of the range of temperatures encountered were used, and both estimates compared. Walleye pollock egg surveys in Shelikof Strait, Alsaka, used temperature at 200m, the maximum depth sampled, in stage length calculations (Kim & Gunderson, 1989), despite variability in egg buoyancy with age. Sundby & Solemdal (1984) used temperature in the homogeneous 0-50m range, where most cod eggs were concentrated.

Sampling variability in ichthyoplankton surveys, due to the non-homogeneous distribution of eggs within a single sampling square and the inevitable randomness of the sampling procedure, has seldomly been directly quantified. A preliminary indication might be provided by an exercise in repeated sampling of a single sampling square carried out during ICES mackerel survey in 1998 (C.Hammer, *pers. com.*) As observed in Table 2.4 the variability observed is in the same magnitude as that uncovered by the present analysis.

Table 2.4: Egg counts obtained during a number of repeated samples undertaken inside the same sampling square (ICES 29D9) in a period of 22 hours (14-15/April/2001).

haul no.	eggs
207.00	318.00
208.00	605.00
209.00	632.00
210.00	566.00
211.00	322.00
212.00	182.00
213.00	230.00
214.00	526.00
mean	422.63
sd	179.05
cv	42%

Finally, uncertainty is also introduced in the estimate by a number of parameters, assumed to be estimated perfectly, but whose point estimates do not reflect the variability present in their estimation. The Bayesian AEPM estimator developed incorporates both error and bias into a coherent framework, providing a measure of confidence around the estimate. It also allows for investigation of the influence in both precision and accuracy of AEP estimates of different survey designs.

The precise bias results shown, based on the 1998 ICES mackerel survey, depend heavily on the temporal patterns of sampling and egg production in that year. The proportion of eggs spawned on both sides of the 200m isobath, for example, has changed over the years (Table 2.3), as has the temporal pattern of spawning. A combination of large spawning in deep waters and early peak in spawning activity, could mean significantly biased DEP estimates being preponderant in the final AEP estimate. Both relevant sources of bias have been quantified by using the only available dataset of mackerel egg distribution with depth from the Northeast Atlantic (Coombs *et al.*, 2001). The question of how representative the data are cannot be easily answered, as no similar sampling of *S. scombrus* egg distribution has taken place elsewhere. Despite the possible limitation imposed by the possible shortages of the egg depth distribution dataset, and its limited coverage, especially in time, the analyses presented provide a reasonable and coherent measure of the bias introduced by two key assumptions in the AEP model, and the variability in results they introduce.

The two sources of bias identified in the ICES mackerel AEP estimate are not limited to this survey. The same assumptions are being made in other implementations of the method, and they will need to be investigated in each individual case before SSB estimates from ichthyoplankton surveys can be used in management decisions with confidence. The distribution of eggs with depth, changes in vertical position with time or age, and differences in temperature along the water column, all need to be considered before decisions on survey design and egg development assumptions are made.

After reviewing each of the variables and processes that conform the AEP method, attention is now turned to the question of spatial and temporal interpolation. Samples taken and processed as explained in this chapter cover a necessarily limited section of the spawning area and season. Interpolation and extrapolation are carried out in both the time and space dimensions to generate a complete estimate of egg production. Geostatistical methods allow detailed modelling of the space and time covariance structure of egg abundance data, and this will be the subject of the next chapter. Variogram models will then be applied to the task of estimating total egg abundance through kriging.

Chapter 3

Bayesian variogram modelling

An important characteristic of ichthyoplankton surveys is their spatio-temporal coverage. In comparison with most stock assessment methods based on catch and effort data from an aggregated unit, egg surveys need to incorporate efficiently both the spatial and temporal dimensions. As explained in Section 2.3, the final result of each individual tow is an estimate of egg density, later to be extrapolated to a certain area of sea surface (in itself a simplification of sea water volume). In the spatial dimension, densities are interpolated at unsampled rectangles. In the temporal one, data first aggregated into points denoting the central day of each of the sampling periods, are then interpolated in time to form a curve of annual egg production. Interpolation and extrapolation in time and space, specially in the context of imperfect sampling, is thus a central characteristic of egg production methods, and could be markedly improved by the use of spatio-temporal statistics. These would allow for explicit modelling of the relationships between samples in time and space, and the use of this information for more precise interpolation in unsampled areas and dates. Furthermore, they provide robust methods for the estimation of the uncertainty in the final egg production estimates.

A Bayesian approach to variogram modelling is presented in this chapter, which allows for untangling of the spatio-temporal covariance structure in the data, its use for point and global estimation through kriging (Chapter 4), and for further investigation of the differences and common trends among subsequent triannual surveys through the use of hierarchical models.

Stock assessment methods based on independent surveys, including the *traditional* implementation of the Annual Egg Production Method, attempt to estimate stock biomass with each new set of data, as if previous years information was simply superseded. Despite environmental or human-driven variations, fish spawning and reproduction is a process linked to a repeating series of factors, and knowledge about its extent and dynamic is cumulative, albeit in many opportunities not formally so. The Bayesian paradigm provides for an elegant and coherent approach for situations where data from repeated instances of the same process, in space or time, is present. Bayesian hierarchical analysis attempts to

estimate the probability distributions describing the possible underlying commonalities among those instances (Gelman *et al.*, 1995). Estimates based on scarce or problematic data can, for example, be improved by what has been learned from previous years. A measure of variability in the parameter estimates describing the process also provides an indication of how variable or similar those processes are across years.

The procedure presented here starts with the estimation of combined experimental variograms in space and time. Data are first transformed to better approximate a Gaussian distribution. A range of variograms models are then fitted in a Bayesian hierarchical context Gelman *et al.* (1995), and a number of convergence and fitting diagnostics are presented. Model selection is then carried out based on the marginal posterior probabilities computed for each alternative model. The benefits and problems of the whole approach are then briefly discussed.

3.1 Estimation of experimental variograms

The variogram, in any of its multiple incarnations, is in essence a function that summarises the degree of similarity or otherwise of data points separated by a certain distance. The basis for its estimation is the calculation of the covariance among all possible combinations of two points in the dataset. Estimation of the experimental variogram and its subsequent modelling is a crucial step in geostatistics (Cressie, 1991). Experimental variograms for mackerel egg abundances were estimated after transforming the data to better approximate the Gaussanity requirements of most geostatistical methods.

3.1.1 Data manipulation

Most geostatistical methods assume the error distribution of the data to be Gaussian, and this is also the case for the variogram modelling and kriging techniques applied in this study. Ichthyoplankton data do not commonly follow a normal distribution, and could be quite often better described as negative binomial. Transformation of data is commonly used in geostatistical analysis, and other fields, to comply with, or to less severely violate, Gaussianity assumptions (Kitanidis & Shen, 1996). More than reflecting on whether the sampled process is in fact multi-Gaussian, the approach applied here is based on the assumption that low-order statistics of Gaussian processes are usually insensitive to moderate deviations from the multi-Gaussian assumption (Cressie, 1991). Like other variance-type statistics, the values of the semivariogram are sensitive to extreme data values (Goovaerts, 1997), but their undue influence can be easily eliminated by a log or power transformation.

Two possible transformations of egg abundance data were considered: logarithmic and Box-Cox transformations (Box & Tiao, 1973; Kitanidis & Shen, 1996). The first one is widely used and is generally well suited to situations with skewed distributions. In this

instance egg densities (m) were transformed as $m^* = \ln(m+1)$. The Box-Cox transform is a power transformation, defined only for positive values,

$$m^*(\lambda) = \begin{cases} \frac{(x^\lambda - 1)}{\lambda}, & \lambda \neq 0 \\ \ln(x), & \lambda = 0 \end{cases} \tag{3.1}$$

where λ, the power factor, is estimated as the value that maximises the logarithm of the likelihood function

$$f(x,\lambda) = -\frac{n}{2} \ln \left[\sum_{i=1}^{n} \frac{(x_i(\lambda) - \bar{x}(\lambda))^2}{n} \right] + (\lambda - 1) \sum_{i=1}^{n} \ln x_i \tag{3.2}$$

and $\bar{x}(\lambda)$ is the arithmetic mean of the transformed data. This transformation has been suggested and applied by some authors for geostatistical data (Ribeiro & Diggle, 2000). A comparison between both methods, showed in most cases no significant differences in

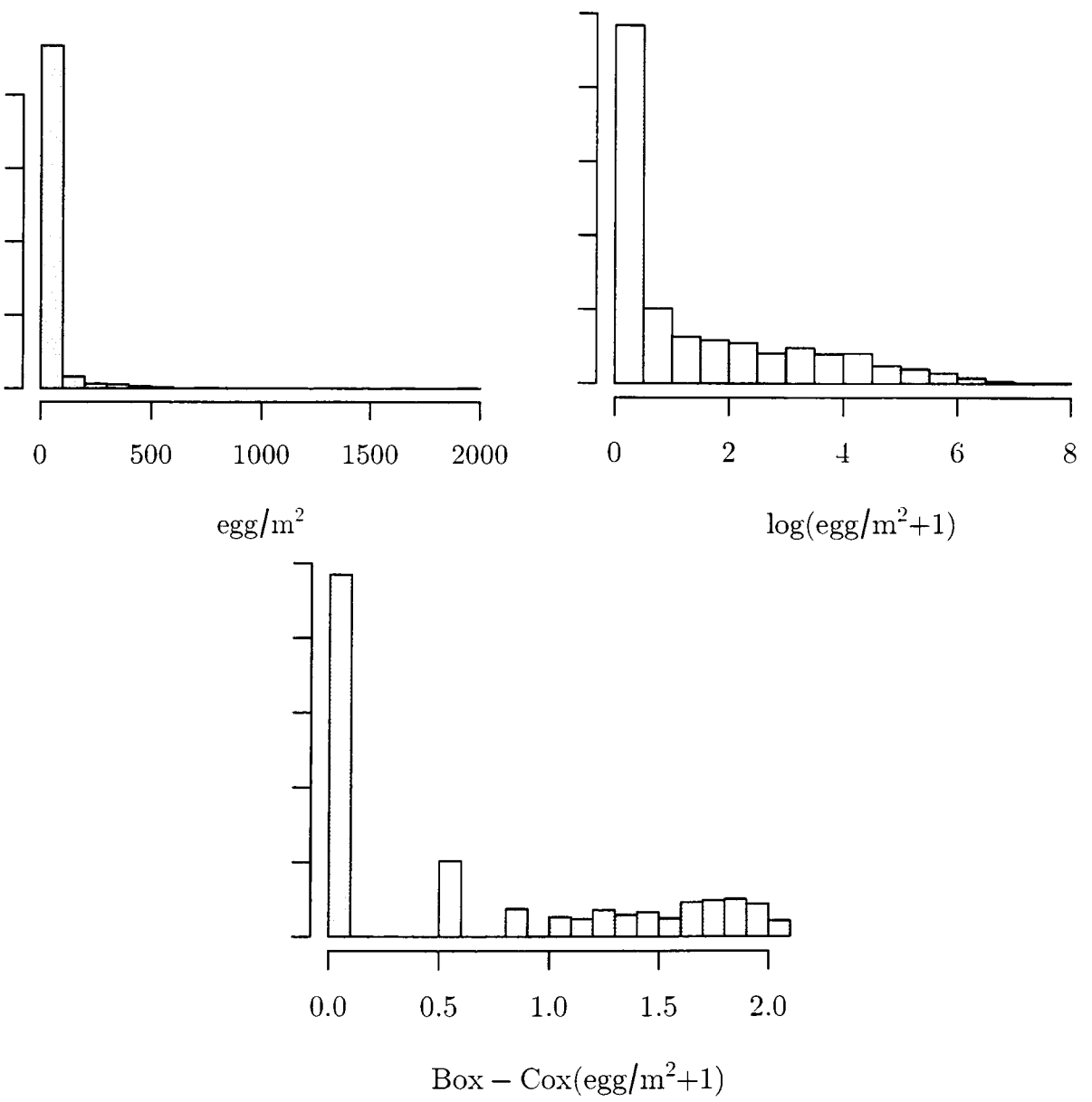


Figure 3.1: Results of the data transformation for the 1998 mackerel egg density (in egg/m^2) using the natural logarithm and the Box-Cox transformation ($\lambda = -0.469$).

the transformed data probability distributions. Differences between the two methods were tested through a chi-square test of goodness of fit (Zar. 1996). An example of both transformations is shown in Figure 3.1 for 1998. Transformed data appears to approximate more closely the desired normal distribution. However, the large numbers of zero values makes perfect transformation impossible, but the transform does help to satisfy the requirements of kriging. Due to its simplicity, both in its application and in the back-transformation of the estimates obtained, the log transform was thus chosen.

The location of each sample in the original datasets is recorded in terms of latitude and longitude. Geostatistical methods depend on Euclidean distances, so positions were transformed into distances in nautical miles from a (0,0) origin set at 60° N - 20° W. This was chosen to lie beyond the most Northwesterly sampled point, recorded at 59° N and 17° W. Calculations used the great-circle distance method (Steinhaus, 1999).

3.1.2 Sample semi-variogram estimation

Of the various measures of spatial continuity and variability used in the geostatistics literature, the semi-variogram is one of the most commonly applied. In contrast with the covariance and correlation functions, the experimental semi-variogram measures the average dissimilarity between data separated by a distance h (Goovaerts, 1997). It is defined as (Matheron, 1965, in Cressie, 1991)

$$\gamma(h) = \frac{1}{2}E\left\{[Z(x+h) - Z(x)]^2\right\} \tag{3·3}$$

where $Z(x)$ is the value of the random function Z at point x. It is computed by the method of moments as half the average squared difference between the components of each data pair whose distance equals h,

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [Z(x+h) - Z(x)]^2 \tag{3·4}$$

where $N(h)$ is the number of data pairs separated by distance h.

where $N(h)$ is the number of pairs separated by distance h. When data are not regularly spaced, the semivariogram estimator is smoothed by defining some tolerance values t around each distance h (Cressie, 1991).

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{\alpha=1}^{N(h)} [Z(x+h\pm t) - Z(x)]^2 \tag{3·5}$$

Spatial phenomena can be anisotropic or isotropic, depending on whether the pattern of spatial variability changes with direction or not (Goovaerts, 1997). Anisotropy is modelled by using directional variograms, in which the value of $\gamma(h)$ depends on the vector $\mathbf{h}$ rather than on the distance $h = |\mathbf{h}|$ only.

3.1.3 Three-dimensional space-time variograms

Interest in this study lies on the estimation and modelling of the complete spatio-temporal structure of egg production. A relatively simple way of applying variography to model spatio-temporal processes is by means of a space-time separable model (although not strictly according to the classification of Kyriakidis & Journel (1999)). The spatial and temporal dimensions of the process under scrutiny are considered to differ in the scale of intensity at which they operate over various distances (the sill), and the maximum distance at which they exert an effect (range). Time is then modelled as a third spatial dimension, such as depth or height would be. Separate spatial and temporal sample variograms are estimated, 2-d and 1-d respectively. Although a number of possible shortcomings with this approach have been identified (Kyriakidis & Journel, 1999), they do not appear to completely invalidate the procedure. Alternatives are being developed (Christakos, 2000), but they still have not being widely applied and are significantly more complex.

The case of the distribution of eggs of a migratory species, such as Atlantic mackerel, is especially amenable to this treatment. The underlying process for the observed spatial and temporal distribution of mackerel eggs is a single one: the northerly migration of the adult stock. What would then be modelled differently is the distinct expression of the same process in time and space. The space-time separable variogram applied in this case includes a omni-directional spatial variogram plus a temporal variogram estimated as corresponding to the vertical dimension of the space-time cube.

3.2 Variogram models

Estimation based on geostatistical methods needs to generate a value for the variable of interest at unsampled points, based on the spatial covariance structure revealed by the variogram. Sample variogram estimators are also not necessarily positive definite (Cressie, 1991), as required by kriging. A valid model (see below) must thus be fitted to the sample variogram. Although a potentially large number of possible variogram models are available, the search was reduced to a reduced number of commonly used and well known models, and to linear combinations of those.

Variogram models can be classified according to their behaviour at the origin and at infinity. All models presented here are bounded, that is, they reach the sill (c) at a distance a (range). Behaviour at the origin can be linear, discontinuous or parabolic.

3.2.1 Valid models

Semivariogram and covariance models must satisfy some conditions if they are to be used in spatial prediction. These conditions help to ensure that no negative variances are produced. The basic requirement is that permissible models must be positive definite (Cressie, 1991). A number of methods exist for testing model validity, but a common approach to avoid such calculations is to use known permissible models and linear combinations of them (Goovaerts, 1997).

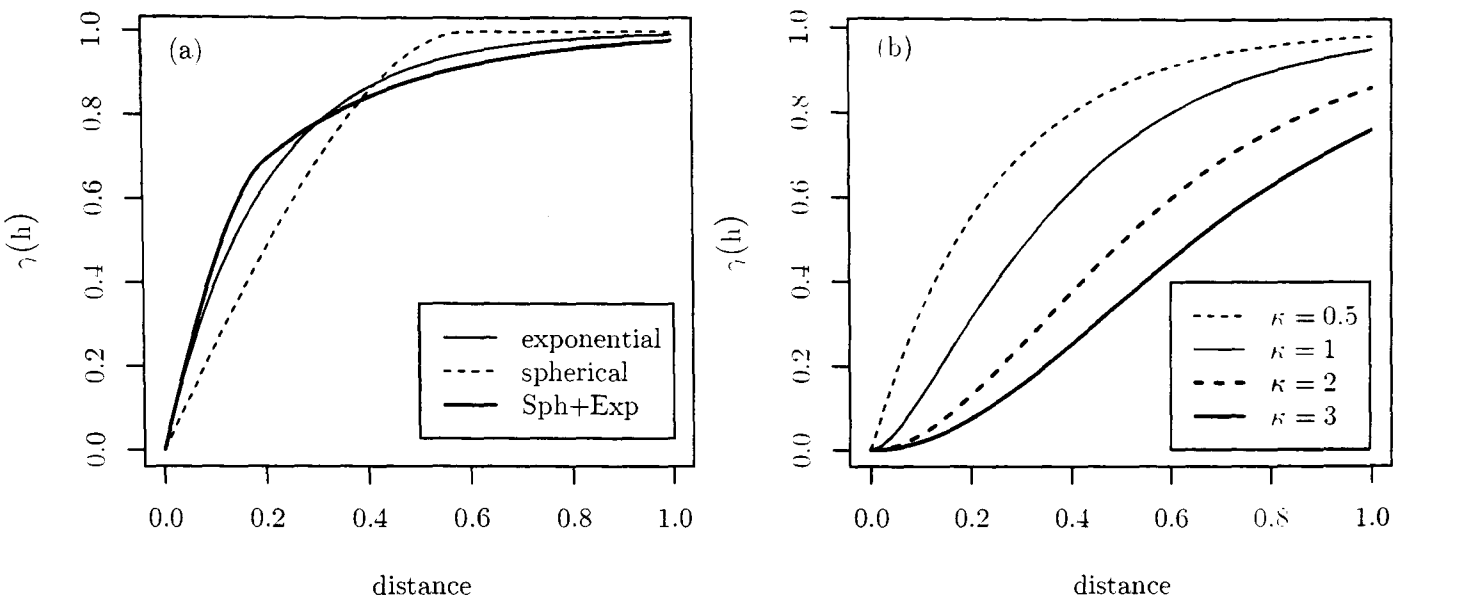


Figure 3.2: Variogram models applied in this study: exponential, spherical and nested combination of both (a), and Matérn (b). The Matérn model in panel *b* is shown for a number of values of the κ shape parameter and for the same combination of range and sill.

3.2.2 Nugget model

The nugget effect model is expressed as

$$\gamma(h) = \begin{cases} 0 & \text{if} h = 0 \\ C & \text{otherwise} \end{cases} \tag{3.6}$$

The sill is reached here as soon as $h > 0$. The nugget is commonly used to model the very short distance variability, due for example to sampling error. This model is clearly discontinuous at the origin.

3.2.3 Exponential model

The exponential model is defined as

$$\gamma(h) = 1 - exp\left(\frac{-3h}{\phi}\right) \tag{3.7}$$

and reaches its sill asymptotically. A practical range is then defined as the distance at which the model value is at 95% of the sill, C (Goovaerts, 1997). The exponential model is linear at the origin, and increases in value faster than the spherical.

3.2.4 Spherical model

The spherical model is defined as

$$\gamma(h) = \begin{cases} 1.5\frac{h}{\phi} - 0.5\left(\frac{h}{\phi}\right)^3 & \text{if } h \leq \phi \\ 1 & \text{otherwise} \end{cases} \tag{3.8}$$

and reaches its sill at the actual range ϕ. This model is also linear at the origin.

3.2.5 The Matérn model

The Matérn model (Mátern, 1986) as presented by Ribeiro & Diggle (2000), is of the form

$$\gamma(h) = \frac{1}{2^{\kappa-1}\Gamma(\kappa)} \left(\frac{h}{\phi}\right)^{\kappa} K_{\kappa}\left(\frac{h}{\phi}\right) \tag{3.9}$$

where $K_\kappa(h)$ is the modified Bessel function of the third kind of order κ, and Γ is the gamma function. The model presents an extra parameter, κ, which controls the smoothness of the random field. The theoretical reasons for its desirability are well covered in Stein (1999), including a critique of functions which do not allow for its smoothness to be estimated (such as the two previously presented), but that lies beyond the scope of the present study. For the purposes of model fitting to sample variograms of egg abundance, its main advantage lies in its flexibility to accommodate large variations in short and medium-range shape. This could prove specially advantageous when all years of data are fitted to a hierarchical model.

At very short distances the Matérn function is parabolic in shape. Thus it should ideally be applied to continuous phenomena. This characteristic is also likely to increase the importance of the *screen effect*, where samples lie between other samples and the point being estimated, often resulting in negative kriging weights (Isaaks & Srivastiva, 1989). Although kriging decreases the weight assigned to screened points, variograms with parabolic behaviour near the origin can lead to negative kriging estimates, often an undesirable outcome. Block kriging, where numerous points at close distances are estimated, is likely to be affected more significantly.

3.2.6 Nested models

Single functions cannot always accommodate the range of variograms encountered. Linear combinations of two or more variogram models can then be applied. These could be of the same functional form, with different parameter values, or a combination of different models. Nested models provide greater flexibility when modelling processes with different behaviours at short, medium or long ranges. All models presented in this study are in fact nested, as they consist of a linear combination of a nugget effect model with any of the others.

3.3 Bayesian hierarchical modelling for spatio-temporal variograms

Estimation of the variogram is usually regarded as the most critical step in the geostatistical procedure (Cressie, 1991), and thus deserves the necessary attention. More so in situations where data are skewed, might not completely conform with the initial model assumptions, or sampling was not designed specifically with geostatistical methods in mind. Variogram fitting has traditionally been a matter of expertise. Models are fitted interactively to the sample variogram by the analyst, maybe using some measurement of

goodness-of-fit, but mostly guided by his/her own perception. Previous experience and knowledge of the process that generated the data is thus informally introduced in the fitting procedure.

Models for a certain process in separate populations of the same species, or any other situation where separate estimates should be considered to be related to each other, can be elegantly accommodated in a Bayesian hierarchical model (Gelman *et al.*, 1995). Parameter values for each single realisation of the process, in our case spatio-temporal variograms of egg abundance in different seasons (j) of mackerel spawning, θ_j, are viewed as a sample from a common population distribution. This is achieved by setting prior probability distributions for the mean and variance of θ. The use of *hyperparameters* can improve the estimation of individual θ_j when data are scarce or relatively uninformative. Information obtained on the hyperparameter probability distributions guides the estimation of individual parameter values for each realisation. In addition, inspection of the posterior probability distributions of the hyperparameters provides useful information on the variability of the parameter values among the different populations. Non hierarchical models might also overfit the present data, limiting the usefulness of the model outputs when applied to subsequent years or new experiments (Gelman *et al.*, 1995)

A hierarchical model for the egg survey variograms consists of two levels: a first one that includes the semivariogram values ($\gamma(h)$) at increasing distances (h) in time and space, and a second one that covers the various years of data, considered in this manner as realisations of a single process. The existence of more than one level gives rise to the use of *hyperparameters*, which in this case consist of the mean and variance of each of the model parameters across different years. The joint posterior density function of the hierarchical model parameters, range (ϕ), sill (C) and nugget (C_0), in given by

$$\begin{aligned} p(C, C_0, \phi|\gamma(h)) = {} & p(\mu_C)p(\sigma_C)p(\mu_{C_0})p(\sigma_{C_0})p(\mu_\phi)p(\sigma_\phi)p(C|\mu_C, \sigma_C)p(C_0|\mu_{C_0}, \sigma_{C_0}) \\ & \times p(\phi|\mu_\phi, \sigma_\phi) \prod_{i=1}^{n} \left(p(C_i|\mu_{Ci}, \sigma_{Ci})p(C_{0i}|\mu_{C_{0i}}, \sigma_{c_{0i}})p(a_i|\mu_{ai}, \sigma_{ai}) \right) \\ & \times \prod_{j=1}^{m} p(\gamma(h)_{ij}|C_i, C_{0i}, \phi_i, \tau) \quad (3.10) \end{aligned}$$

where $\gamma(h)$ represents the modelled data, the semivariogram values at distances h, and i and j are the number of years and the number of datapoints per year, respectively. τ is the variance in the semivariogram values.

3.3.1 Selection of prior probability distributions

A delicate issue in Bayesian analysis is the question of prior selection (Gelman *et al.*, 1995). As no previous information that could help in constructing informative priors for the variogram parameters in these analyses existed, relatively non-informative priors were chosen. However, completely flat priors were avoided. The range of possible values

for the four variogram parameters under study – sill, nugget, range and, in the case of the Matérn function, κ – had to be strictly positive. A maximum value could also be imposed without reasonably affecting the final result. Relatively flat truncated normal and log-normal probability distributions were then chosen, (See Table 3.1). The possible impact of prior choice was then investigated by repeated runs of the same analyses with different prior distributions specified.

Table 3.1: Prior probability distributions for the variogram parameters. Normal probability distributions are expressed in terms of μ and σ^2.

Parameter	Distribution	Lower limit	Upper limit
sill (c)	Normal$(4, 300^2)$	$1 \cdot 10^{-8}$	$1 \cdot 10^{5}$
nugget (c_0)	Lognormal$(0.5, 50)$	$1 \cdot 10^{-8}$	$1 \cdot 10^{5}$
range (a)	Normal$(100, 500^2)$	$1 \cdot 10^{-8}$	$1 \cdot 10^{5}$
kappa (κ)	Normal$(1, 300^2)$	$1 \cdot 10^{-8}$	$1 \cdot 10^{5}$

3.3.2 Likelihood function

Given the experimental variogram values, X, and some parametric approximation to them, $f(\theta)$, for N data points, the standard normal likelihood function is given by

$$\mathcal{L}(X \mid \theta) = \prod_{n=1}^{N} \frac{1}{\sqrt{2\pi\sigma_X^2}} \exp\left(-\frac{(X_n - f(\theta))^2}{2\sigma_X^2}\right) \tag{3.11}$$

and its associated log-likelihood is given by the following:

$$\ln \mathcal{L}(X \mid \theta) = -\frac{N}{2}\ln(2\pi\sigma_X^2) - \frac{1}{2\sigma_X^2}\sum_{n=1}^{N}(X_n - f(\theta))^2 \tag{3.12}$$

3.3.3 MCMC integration

Bayesian analysis of complex models usually entails integrating over all possible parameter values. The integral on the denominator of Bayes' theorem has commonly no closed form solution and needs to be solved numerically. To achieve this a Markov Chain Monte Carlo (MCMC) (Gamerman, 1997; Brooks, 1998) procedure was applied. A standard Metropolis-Hastings algorithm (Metropolis *et al.*, 1953; Hastings, 1970) was used in this case.

Significantly more efficient MCMC samplers can be achieved by carefully adjusting the algorithm to the problem at hand. For example, convergence can be improved and speeded up by combining a Metropolis-Hastings with a Gibbs sampler (Brooks, 1998). Also, use can be made of conjugates when selecting the functional forms of some priors. Parameters would then be updated conditionally on values for other parameters previously accepted by the sampler, thus both simplifying and speeding up the sampler.

Although tailor-made samplers are in general desirable, they lack in generality, and require additional effort if a range of models are to be explored. As the intention of the present analysis was to explore various functional forms for the variograms, a decision was made to apply a more general MCMC sampler that did not need to be adjusted for each model. This lead to a decrease in efficiency and slower mixing of the chains, but resulted in more flexibility and adaptability to multiple models and parameterisations. The *MCsim* package provided a simple way for implementing the Metropolis-Hastings MCMC samplers both flexibly and efficiently.

3.3.4 Convergence diagnostics

Convergence of Markov chains needs to be carefully assessed. For this a series of diagnostics were applied. First, the chain or chains must proceed long enough, otherwise they could be grossly unrepresentative of the target distribution (Gelman *et al.*, 1995). In the present analyses, as discussed in the presentation of methods, simulations consisted of a minimum of 100,000 iterations. Even when the chains have run for a long period, the early iterations reflect the prior distributions more than the target ones, so they need to be discarded when making inferences. In this case a *burn in* period of 25,000 iterations was implemented, and only the remaining values were used. The most obvious test of chain convergence is to compare the values obtained at its start and end for each of the estimated parameters. If the chain has reached stability, the values should not differ significantly.

Within-sequence correlation can also be problematic, as the assumption of independence of each random draw is violated by the dependency introduced by the Markov chain. This can be assessed and, if present, solved, by *thinning* of the chain at regular intervals. A comparison of the posterior probability distributions obtained from chains thinned at various intervals should again show no significant difference.

Finally, a number of parallel chains with different starting points can be compared. And again they should show no significant differences. The first two tests were applied to all simulations, which were then run for another 100,000 iterations if any of the estimated parameters failed those tests. Simulations with parallel chains were conducted in a limited number of cases, as they imply a significant increase in computing time.

The CODA software (Best *et al.*, 1996) comprises a suite of tools for the assessment of convergence of MCMC chains, programmed in the *R* statistical language (R Development Core Team, 2004). Although initially devised for use with chains output by Gibbs sampling, most methods included are general enough to be applied on results of Metropolis-Hastings MCMC runs. The convergence diagnostics implemented include a number of visual procedures, plus formal tests such as those proposed by Gelman & Rubin (1992), Raftery & Lewis (1992), Heidelberger & Welch (1983), and Geweke (1992). Chains were assessed using these tests, and rerun for an extra 100,000 iterations and re-tested whenever they showed lack of convergence.

3.3.5 Bayesian model checking and sensitivity analysis

Model fitting was assessed by estimating Bayesian, or posterior predictive, p-values (Gelman *et al.*, 1995). Simulations were carried out using 1000 draws extracted from the last 5000 at the tail of the MCMC chain. Datapoints were re-estimated using those parameter values. The p-value is defined as the probability of the replicated data being more extreme than the observed data,

$$p - value = P\left(\mid T(y^{rep}, \theta) \mid \geq \mid T(y, \theta) | y \mid \right) \tag{3.13}$$

3.3.6 Use of Bayes' posteriors for model choice

Each of the analyses presented here have as a starting point a single functional model. But choice of functional form for the model variogram is far from clear, and significantly different results could arise if an alternative model was applied. The posterior probability of model i can be approximately estimated as

$$P(M_i|d) = \frac{P(d|M_i)P(M_i)}{P(d)} = \frac{P(d|M_i)P(M_i)}{\sum_j P(d|M_j)P(M_j)} \tag{3.14}$$

where d represents the data, in this case semivariogram values (γ) at different distances. By ignoring the normalisation constant, equation 3.14 can be rewritten as

$$P(M_i|d) \propto \int_{\theta_i} P(d|\theta_i, M_i)P(\theta_i|M_i)d\theta P(M_i) \tag{3.15}$$

Integrals of this form can be evaluated through the Monte Carlo method by generating a sample from the density $P^*(\theta)$.

A commonly applied model selection procedure is the Bayesian Information Criterion (Schwarz, 1978), developed as an approximation of the integrated likelihood, and defined for model m as

$$BIC_m = -2 * L_m + v_m \log(n) \tag{3.16}$$

where $L_m = \log f(x|m, \hat{\theta})$ is the maximum log-likelihood for model m, and v_m is the number of estimated parameters in the model.

Estimating the number of free parameters in a hierarchical model is not without difficulty, due to the presence of hyper-parameters (Spiegelhalter *et al.*, 2002). These authors proposed an alternative metric, called *Deviance information criterion* (DIC), as a combination of an effective measure of the number of parameters (p_D) and the posterior mean deviance, defined for the exponential family as

$$D(\theta) = -2 \log\{p(y|\theta)\} + 2 \log f(y) \tag{3.17}$$

where $f(y) = p\{y|\mu(\theta) = y\}$.

3.3.7 Implementation of the Bayesian and geostatistical methods applied

The analyses presented in this chapter have been implemented using a mixture of software platforms. Experimental variograms were estimated using *gstat* (Pebesma & Wesseling, 1998). Bayesian models were implemented in *MCsim*, a free-software library that generates C code for a Metropolis-Hastings MCMC sampler from a model specification in an *R*-like language. The Matérn function was then manually incorporated into the resulting C code before compilation. The need for this last step prevented the use of a common choice, the WinBUGS modelling software (version 1.3), as its closed-source nature and the lack of an extension mechanism did not allow the Matérn function to be implemented within it.

3.4 Results

3.4.1 Sample variograms

The experimental variograms estimated on the log-transformed egg production data are presented in Figure 3.3. The omni-directional spatial variogram, estimated on the xy plane, and the temporal one, on the z plane, are shown together, although with different units on the x-axis (nautical miles and days).

Variograms appear to differ markedly among years. Semivariograms in the spatial plane have small nugget values for most years, that appear noticeably larger for the last three. A local scale of around 60 nm can be distinguished, while behaviour at longer distances is more erratic. Variograms for 1983 to 1989 appear to be dominated by the long-range structure. Unexpected apparent increases in covariance with distance are present in some of them.

In the temporal plane, variogram shapes might be considered less variable, but are still far from consistent. The large nugget value previously observed for 1995 is replicated in this dimension, and is perhaps indication of higher sampling uncertainty. Nugget values are generally lower in the temporal plane. This could be interpreted as arising from the concentration in time of sampling effort. The transition between short and long-range structures is found at between 20 and 40 days.

3.4.2 Assessing the exchangeability assumption

Hierarchical models assume probability distributions of the parameters for each experiment, survey years in this case, to be exchangeable (Gelman *et al.*, 1995). The joint probability distribution of parameter values for each year y, $p(\theta_i, \cdots, \theta_y)$ must be invariant to permutations of the indexes $(1, \cdots, y)$. This assumption is often a reasonable one even if we know processes to vary across years.

A visual inspection of the sample variograms presented in Figure 3.3 shows the existence of two distinct variogram shapes. Some years show a relatively flat or asymptotic variogram while others show sharper linear shapes. It might thus seem that attempting to model the two contrasting variogram shapes within a single model structure might

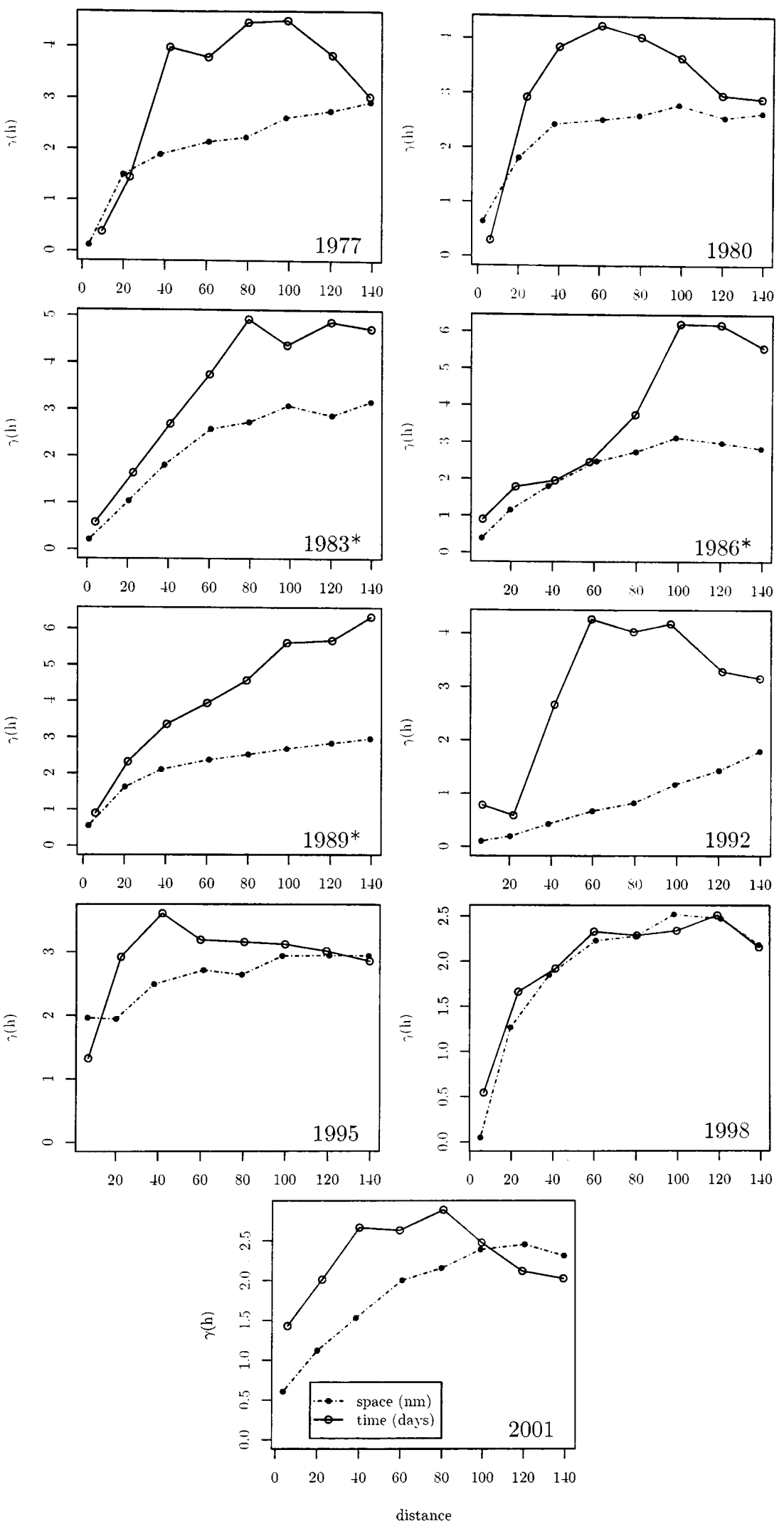


Figure 3.3: Sample semivariograms, $\gamma(h)$, in space (x,y) and time (z) for mackerel egg abundances. h equals the distance, in nautical miles in the space variogram, in days in the temporal one. Plots marked (*) refer to years modelled in a separate group.

lead to unreasonably large uncertainties in the final probability distributions. If the exchangeability assumption does not hold, two procedures are available. Ideally, one should have available information on covariates that explain the difference observed. The probabilities would then become exchangeable conditioned on the covariate. When no physical explanation for the differences exist, as is the case here, the model needs to be split in two groups so that exchangeability can be assumed within each of them. In this way the model gains in predictive power for individual years at the expense of generality.

No clear explanation could be found to account for the two variogram behaviours observed. A decision was taken to separate the survey years in two groups, so called *asymptotic* and *linear*. The first group included 1977, 1980, 1992, 1995, 1998 and 2001, while the second contained 1983, 1986 and 1989. Hierarchical analyses were conducted using all years or separate for both variogram shapes. Results were significantly different, in terms of model residuals, convergence and posterior variance. A decision was taken to continue with the two-model construct for kriging. Those models form the body of results shown in this chapter. Comparisons were later carried out on the differences in kriging results between the one and two-model analyses.

3.4.3 MCMC Convergence

Convergence of most of the MCMC chains was relatively slow, as expected for models with more than a few parameters, and given the known problems of correlation among some of them. Chains had to be run for lengths of between 250,000 and 2,000,000 iterations before convergence could be achieved. Parameter values for one year and in two models still showed poor mixing or significant autocorrelation at quite distant lags. This was never the case with the hyperparameters. Chains showed autocorrelation larger than 0.1 at lags of between 3 and 12, and in each case a thinning interval was applied to ensure autocorrelation to be below that value.

3.4.4 Prior sensitivity analysis

Vague prior probability distributions were applied in the present analysis, only limiting in practical terms the range of possible values (see Table 3.1). Basic checks on the influence of those priors were nevertheless carried out. Informative priors for all mean hyperparameters were defined, with probability densities skewed towards either extreme of the original distributions. Analyses were then repeated. An example of the differences in the obtained posterior probability distributions are presented in Figure 3.5. Results did not vary significantly across the two prior sets, as tested through a χ^2 test (Zar, 1996). This was due to the highly informative data.

3.4.5 Model fits

All models provided reasonable fits to the sample variograms (Figures 3.6-3.9). Some of the model variograms might nevertheless provide very little information when kriging is applied. For example, fits to the temporal variogram by the Matérn model and for 1995

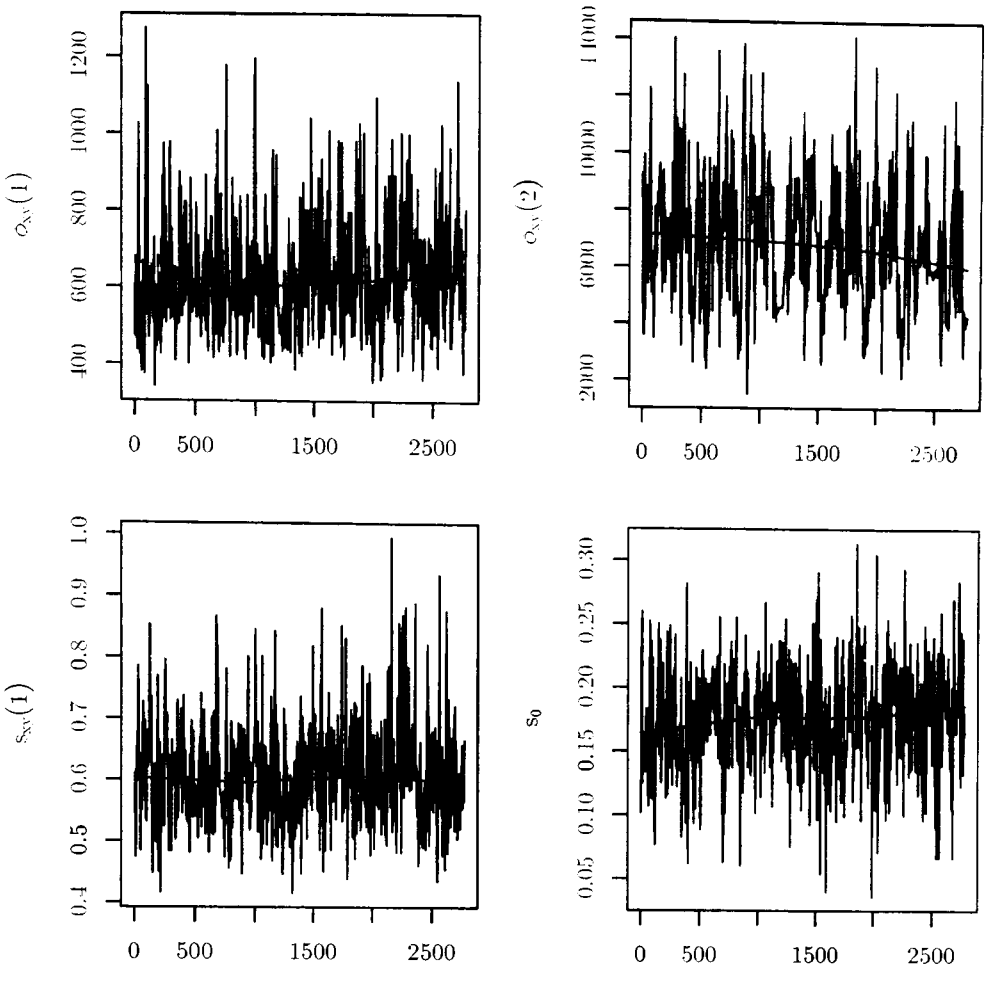


Figure 3.4: Example MCMC chain traces of the nested model hyperparameters for the nested model.

and 2001 appeared to be surprisingly steep and with very short ranges (Figure 3.6), in clear contrast to the other models.

3.4.6 Parameter estimation and uncertainty

The posterior distributions of the various hyperparameters show the range of variation in variogram shapes among years (Table 3.2). The largest variance can be found in the estimates of C_0, with CVs well above those found in the other parameters. A comparison between models should not be carried out directly, specially for both ranges (ϕ_x for the spatial and ϕ_z for the time plane), as they are related to each model's structure. It is worth noting how relatively low the variation in ranges across years appears to be, specially when compared with the variation observed among models.

Table 3.2: Mean and CVs for the different hyperparameters under the four models tested (Matérn, exponential, spherical and exponential + spherical). Simbols as in Appendix A.

	Mat	Exp	Sph	Nes
Spatial range (ϕ_x)	128.68 (19.4)	291.66 (10.3)	227.79 (11.1)	4091.73 (31.4)
Temporal range (ϕ_z)	1.82 (53.5)	147.33 (14.9)	94.08 (15.3)	237.73 (26.8)
Nugget (C_0)	4.12e−04 (148.9)	0.475 (31)	0.715 (105.3)	0.158 (47.8)
Sill (s)	2.12 (15.9)	2.93 (13)	2.4 (13.1)	2.89 (13.3)

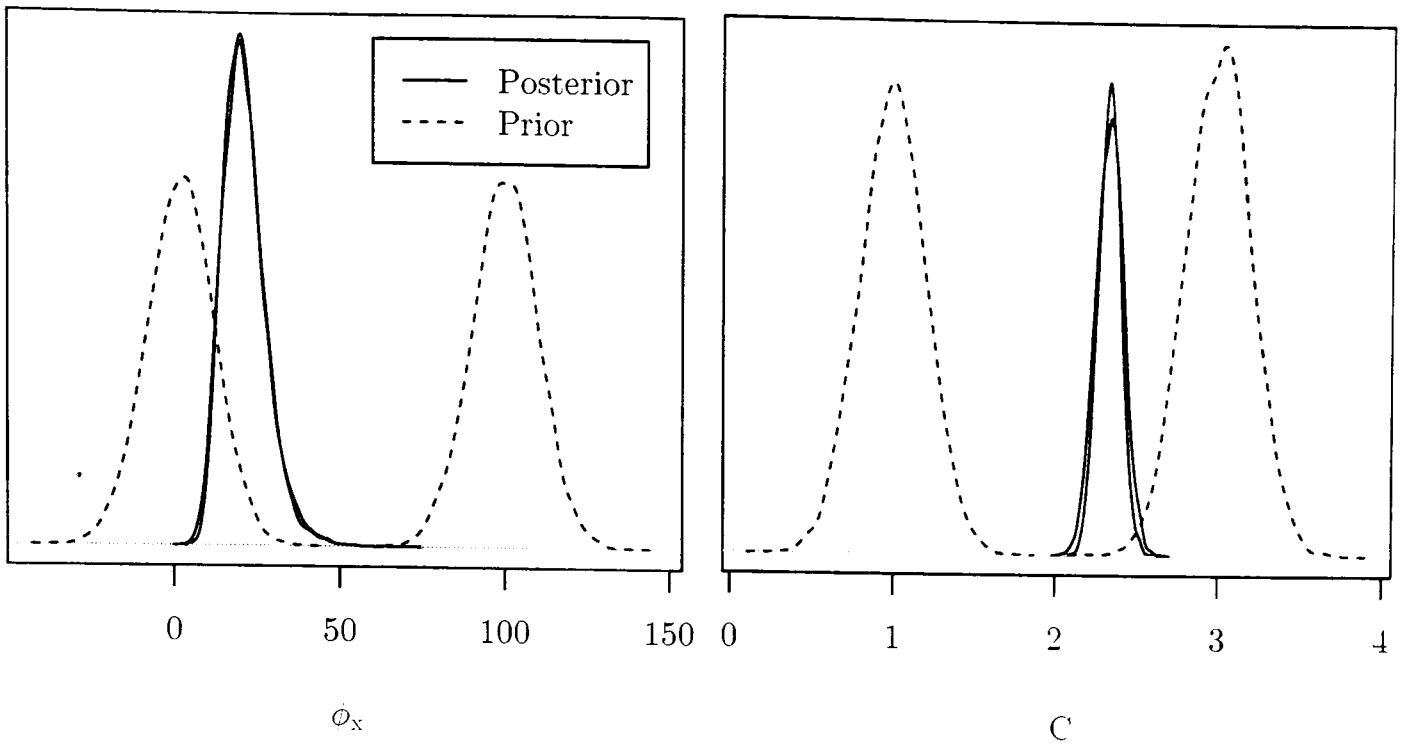


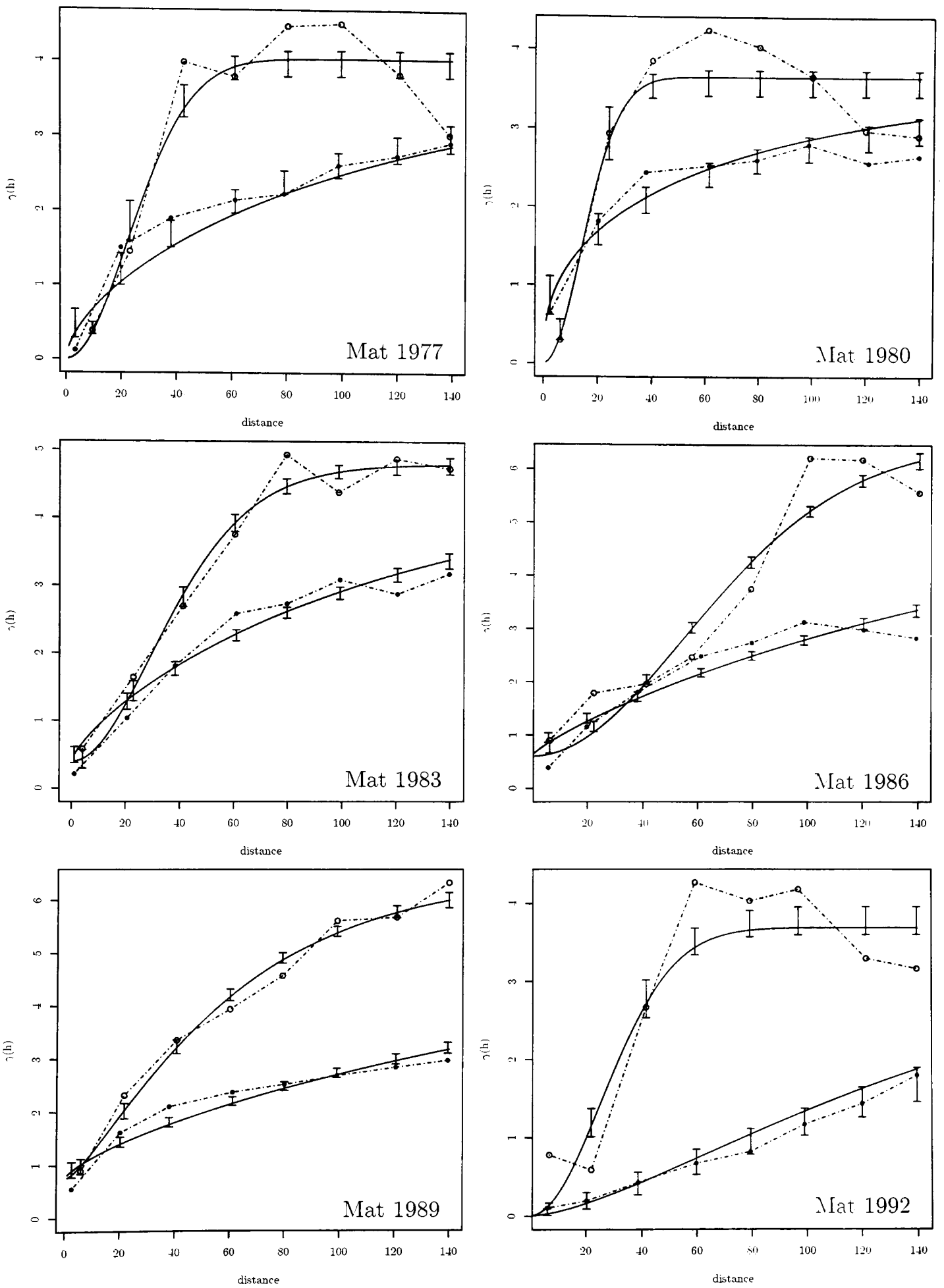
Figure 3.5: Posterior probability distributions (thin line) for ϕ_x and C obtained with divergent prior probability distributions (dotted line).

For some years, the posterior probability of C_0 for the Matérn model was extremely narrow and close to zero (Figure 3.12). The real shape of the pdf for those years where a single line is present can be summarised by the one obtained for 2001 (Figure 3.13). This appears to be related to the flexibility of the Matérn function, that allows it to change in shape and vary in range and sill, while at the same time remain around a limited range of values for the nugget.

The most important parameters influencing the kriging estimates are the (highly correlated) sill and range (Isaaks & Srivastiva, 1989). Plots showing the joint distributions obtained for the combinations of both ϕ_x and ϕ_z against sill are presented in figures 3.14 and 3.15. These plots provide further indication for the separation of the analysed years in two groups. Again the years from 1983 to 1989 show distinct patterns, in this case of higher correlation among range and sill. 1992 highlights the difficulty of separating variograms in this case. Slight correlation can also be observed for the 1995 spatial and temporal parameters (Figures 3.14 and 3.15, bottom left panel).

3.4.7 Relative nugget

The relative nugget is defined as the ratio of the nugget discontinuity to the sill value (Goovaerts, 1997). It allows comparison among years and models of the estimates of sampling error and short distance variation. Mean values for this parameter range between 2 and 10% of the sill value, although in some years the posterior distributions reflect a wide uncertainty (Figure 3.16)


Mat 1977
Mat 1980
Mat 1983
Mat 1986
Mat 1989
Mat 1992
γ(h)
distance

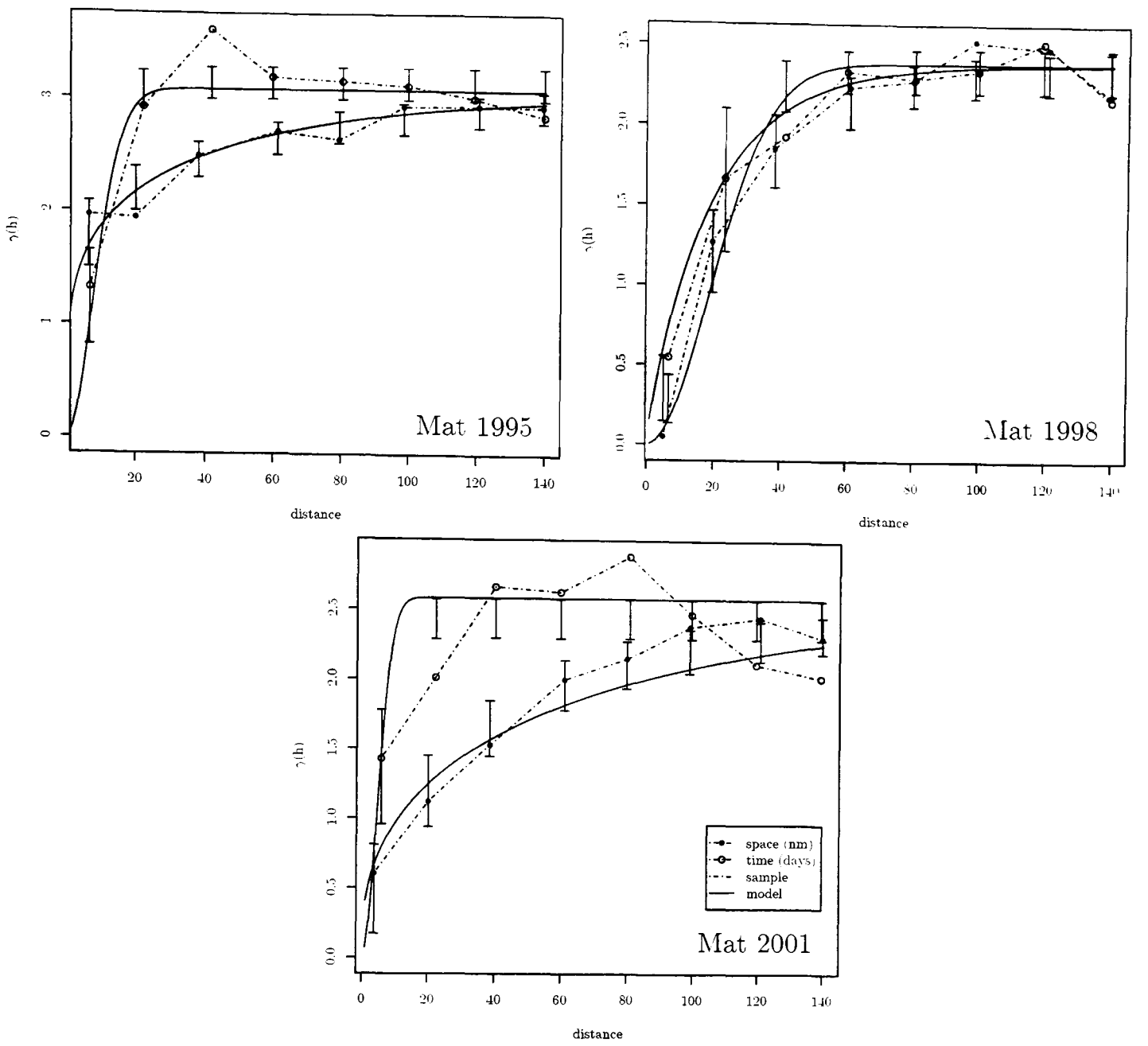
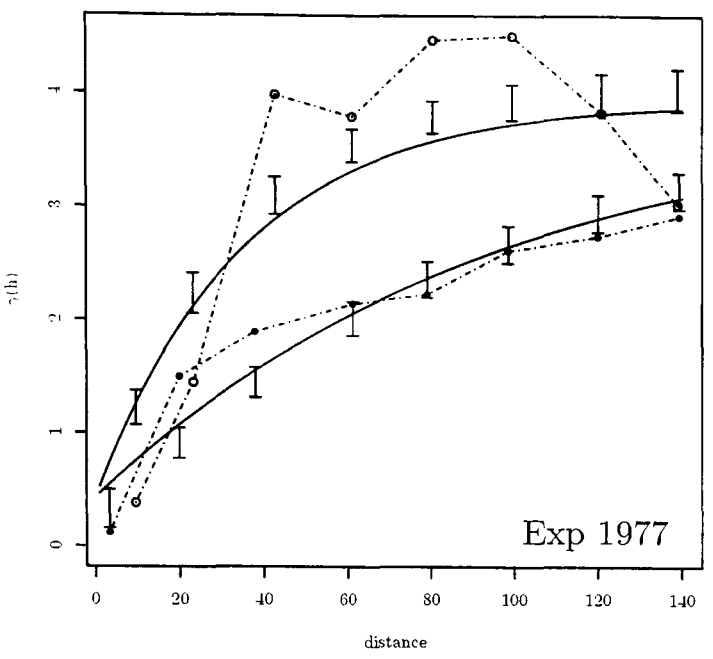
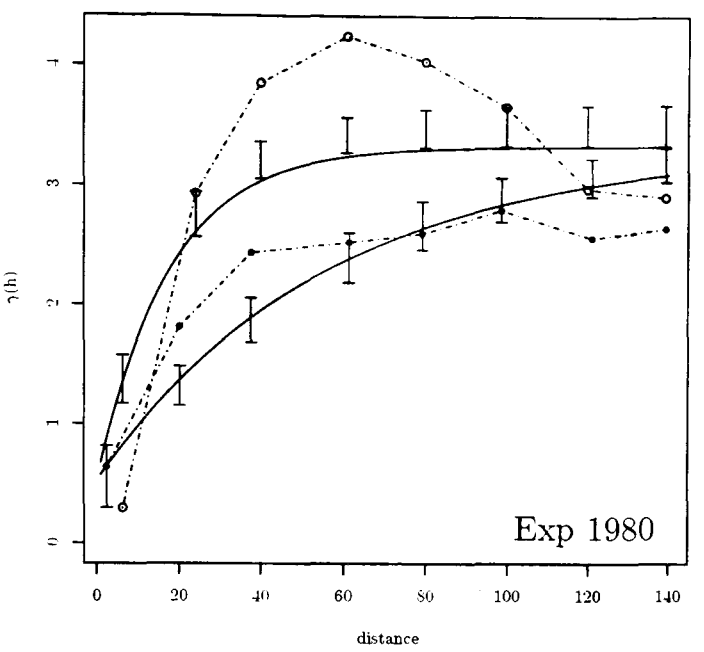


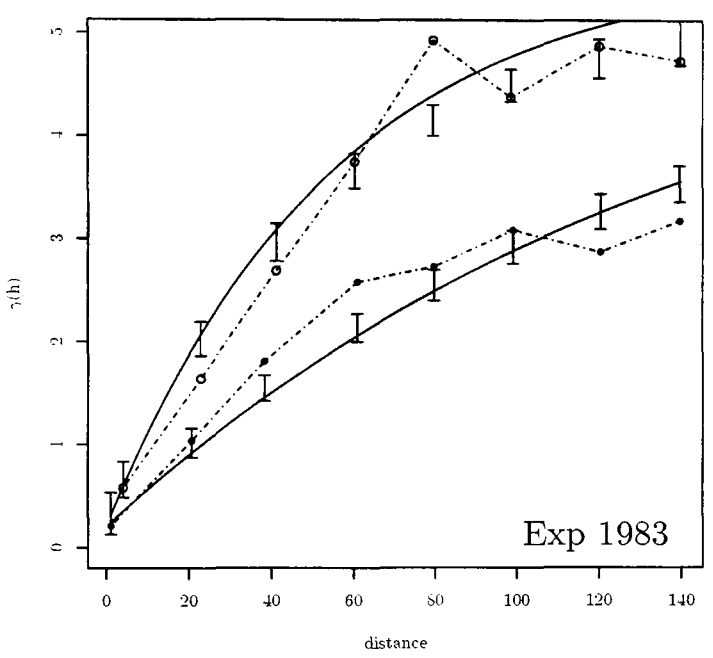
Figure 3.6: Variograms obtained with the Matérn model. Plots show the maximum likelihood model.

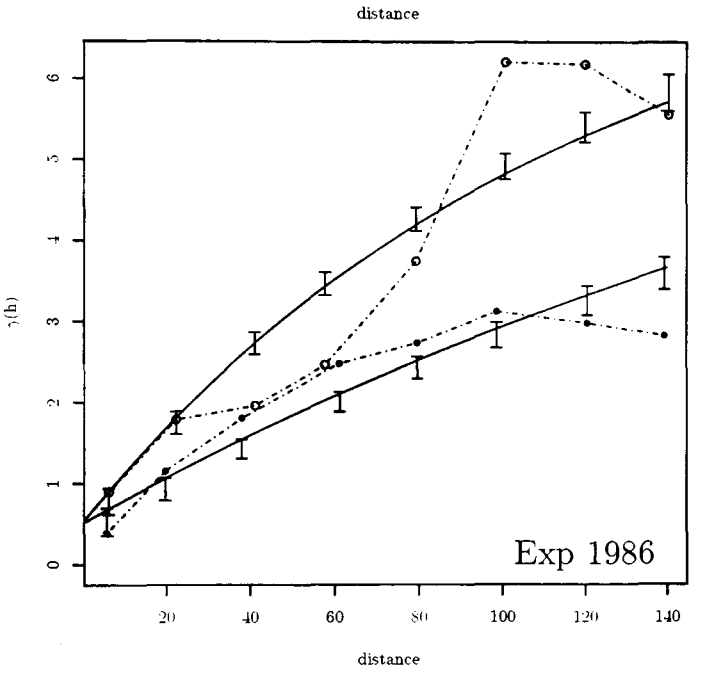
3.4.8 Model choice

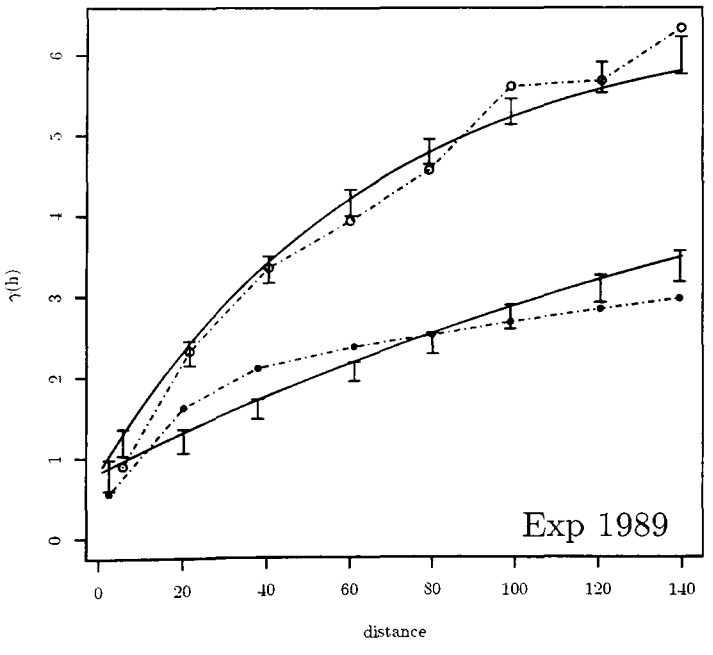
Estimates of the different model complexity and fit metrics can be found in table 3.3. The simpler models appear to fare significantly better. Neither the nested or the Matérn model improve the fit to the sample variograms to such as degree as to compensate for the extra number of parameters to estimate. Calculation of most of the model choice metrics in hierarchical models faces the difficulty of delimiting the actual number of parameters being estimated (Spiegelhalter *et al.*, 2002). No clear solution appears to be available. As the hierarchical structure proposed here is equivalent across all models, the number used in the calculation of both DIC and BIC was simply the number of parameters (5 or 7 depending on the models) plus the number of hyperparamaters (3), without any correction.

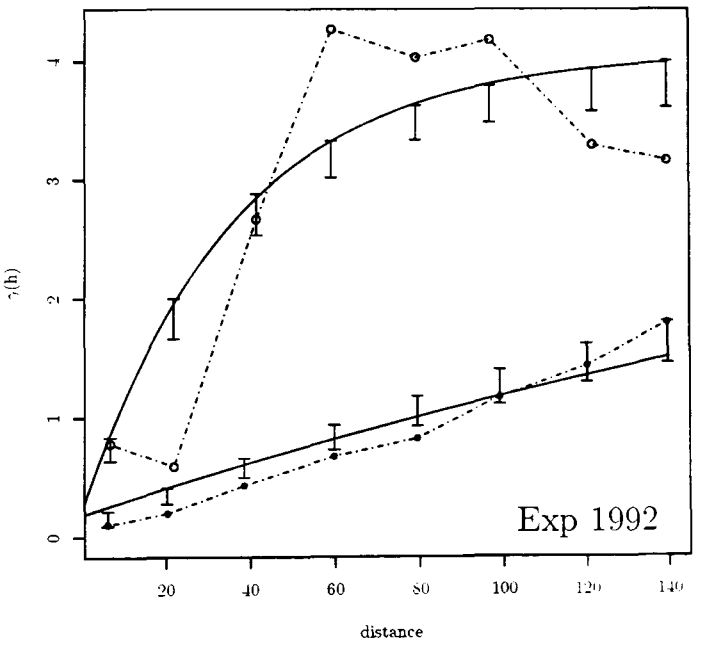

Exp 1977
γ(h)
distance


Exp 1980
γ(h)
distance


Exp 1983
γ(h)
distance


Exp 1986
γ(h)
distance


Exp 1989
γ(h)
distance


Exp 1992
γ(h)
distance

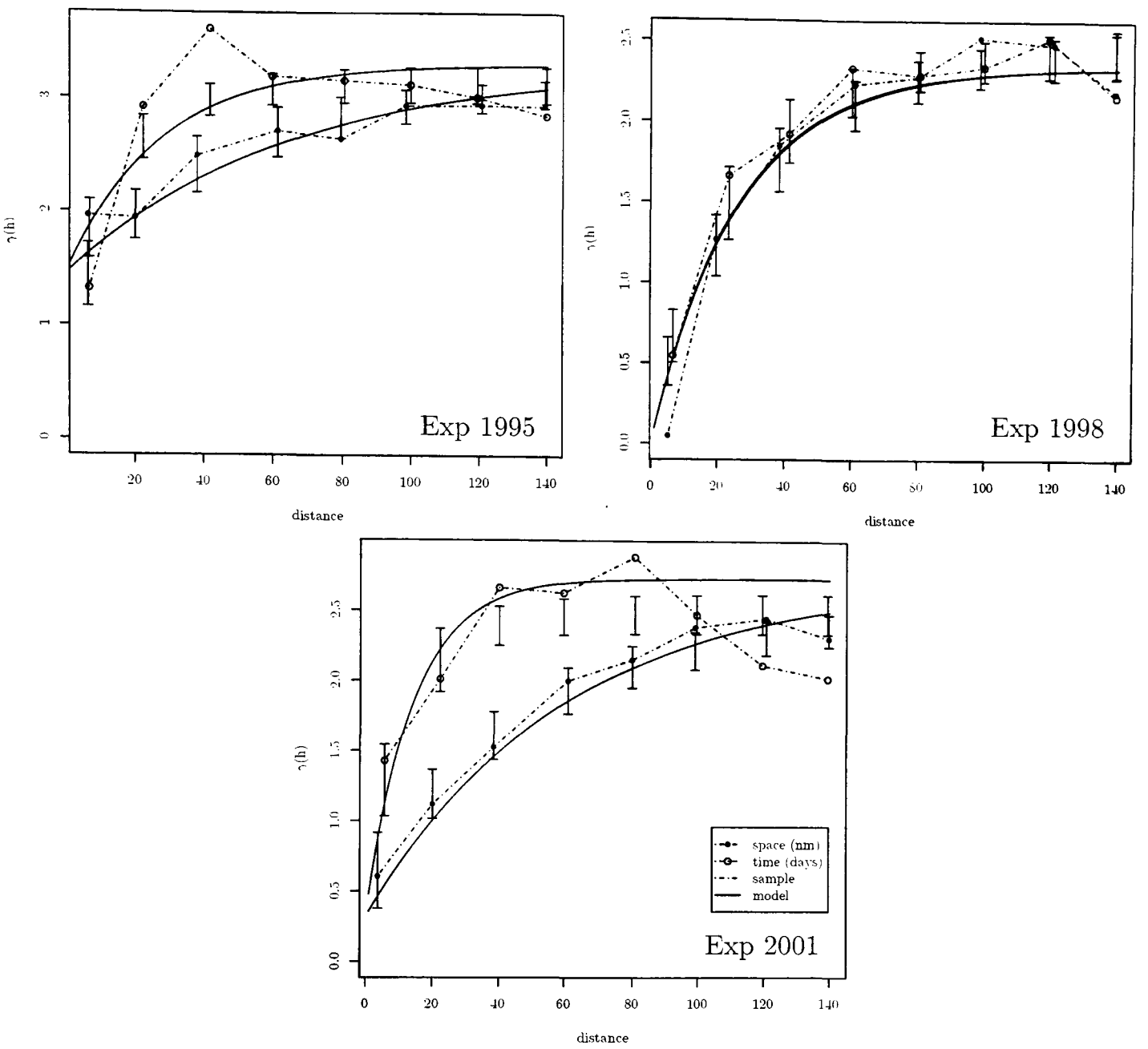
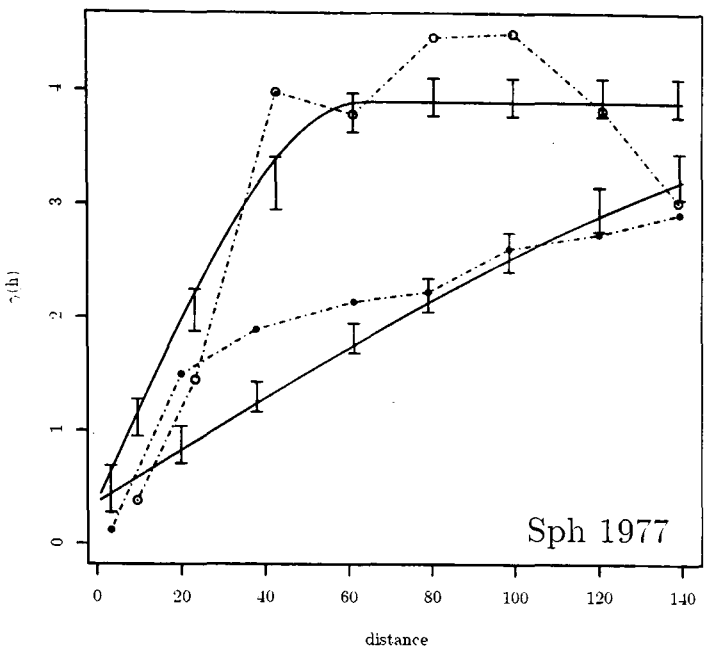
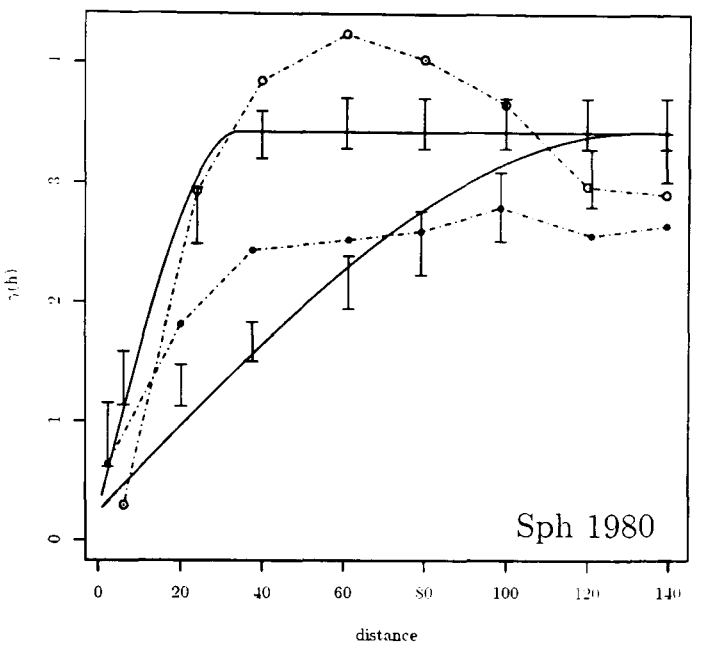


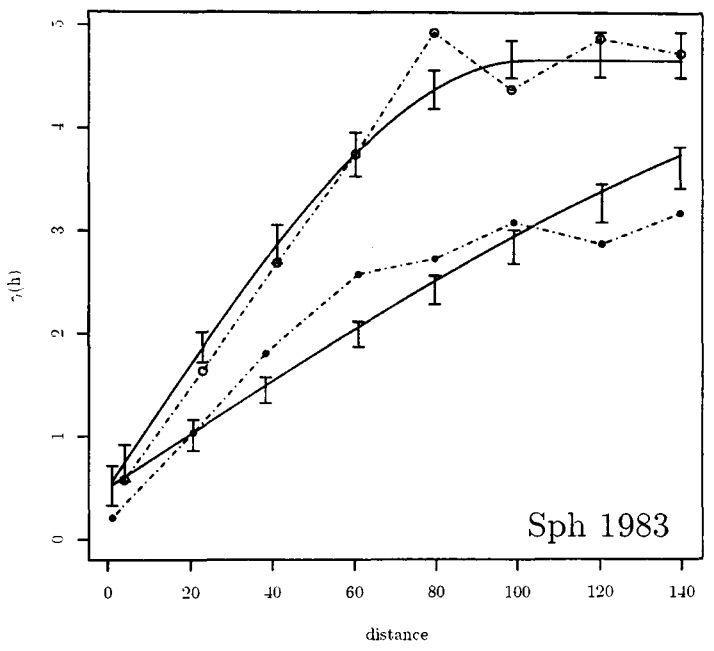
Figure 3.7: Variograms obtained with the exponential model. Plots show the maximum likelihood model and the 95% credibility interval.

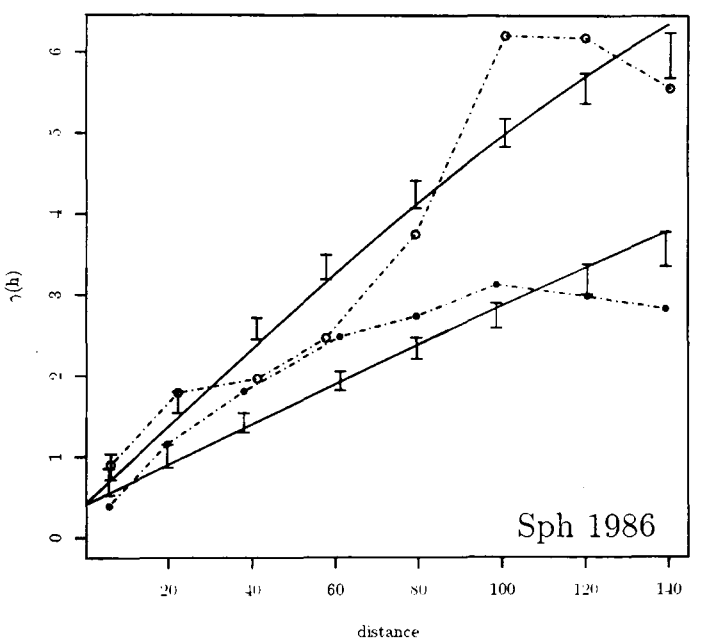
Table 3.3: Values obtained for the various model fit metrics and for each model.

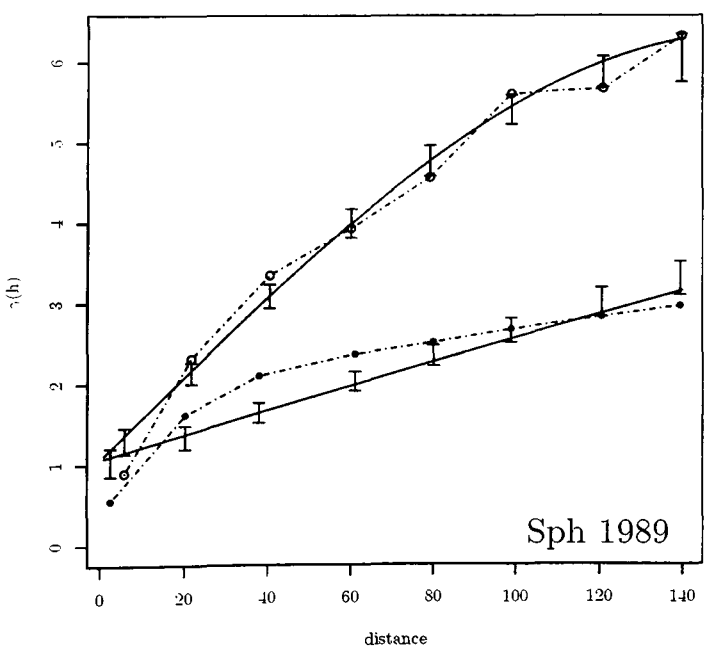
model	Harmonic mean	BIC	DIC	parameters
Matern	9886.3	27.87	29.21	18
Exponential	10910.1	46.42	36.01	12
Spherical	10210.7	45.73	34.22	12
Nested	8928.0	13.06	30.12	21

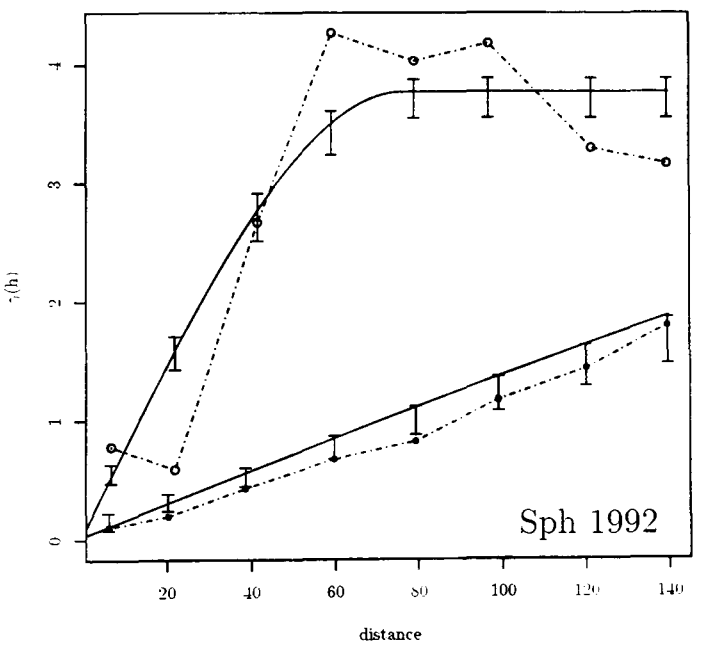

Sph 1977
γ(h)
distance


Sph 1980
γ(h)
distance


Sph 1983
γ(h)
distance


Sph 1986
γ(h)
distance


Sph 1989
γ(h)
distance


Sph 1992
γ(h)
distance

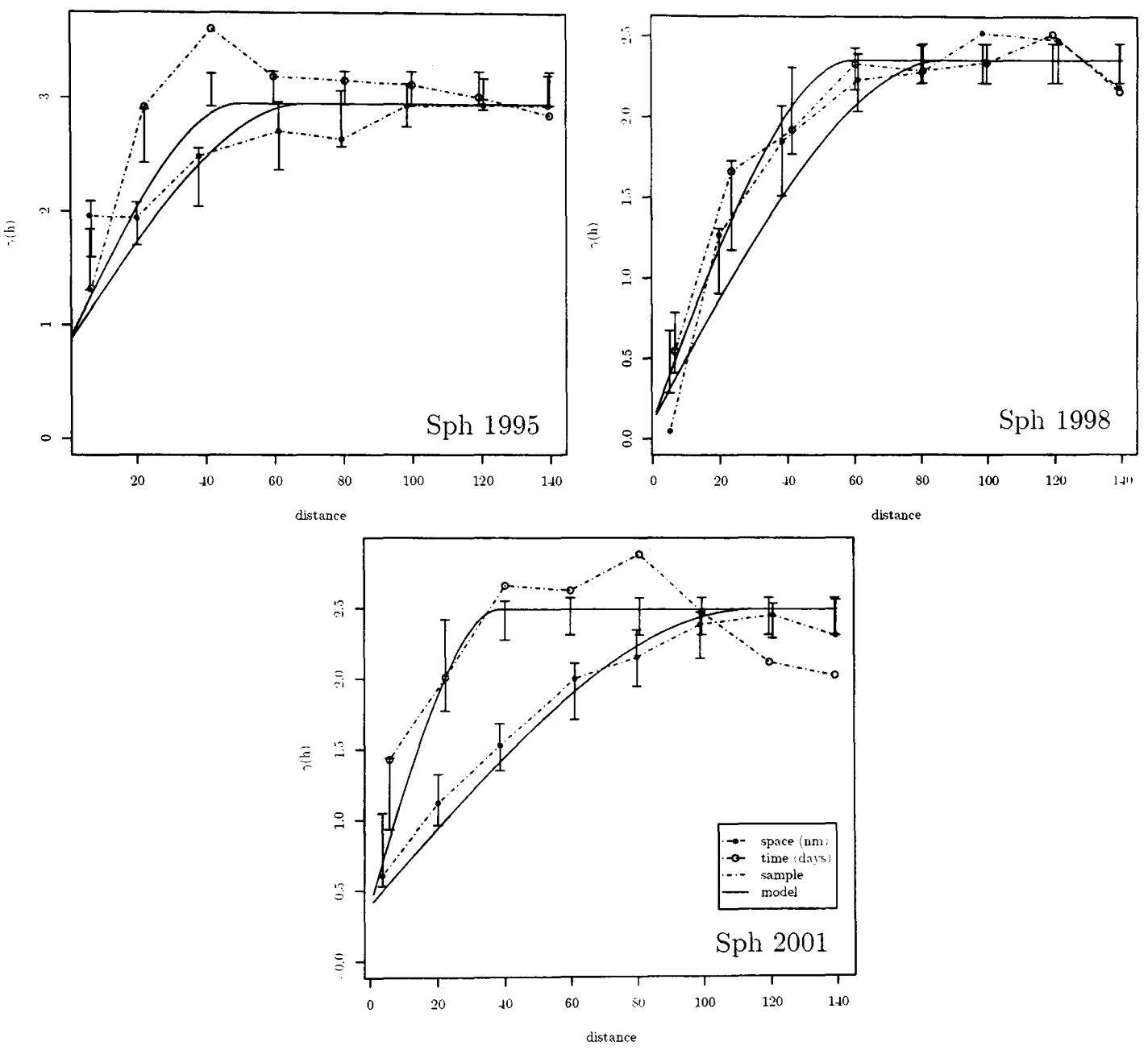
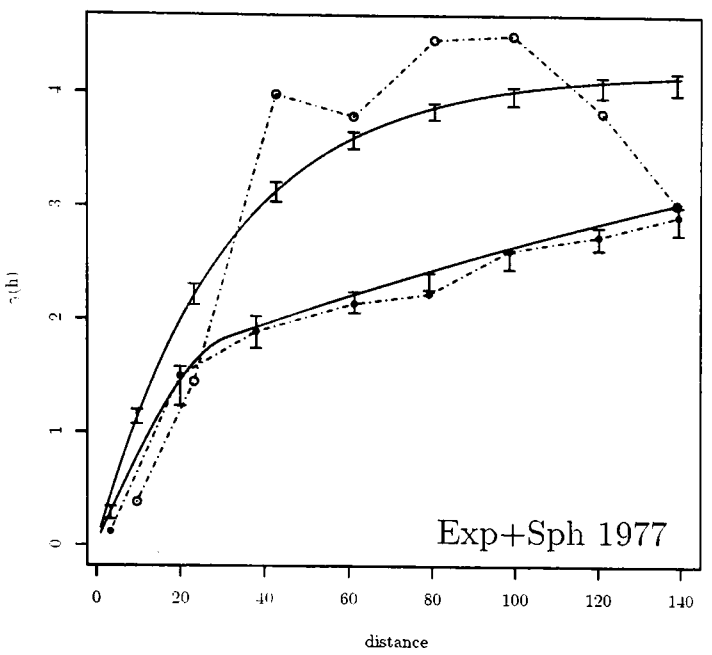
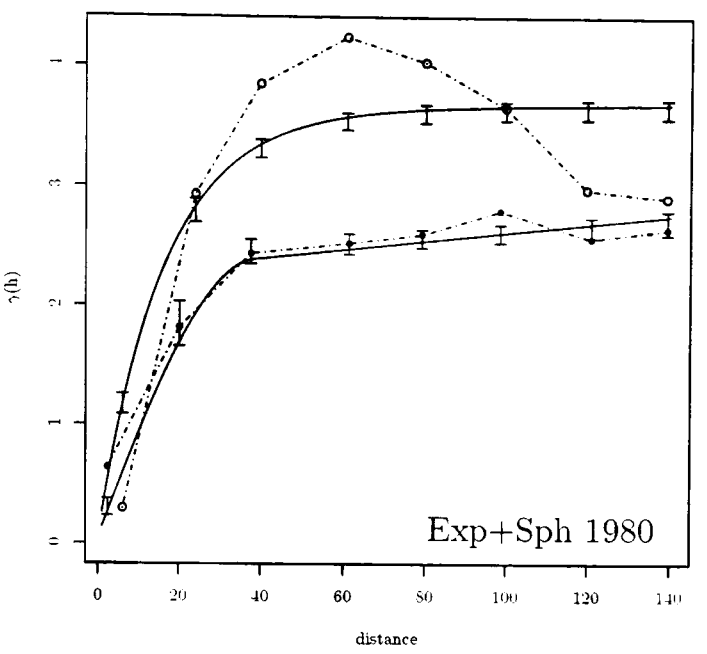
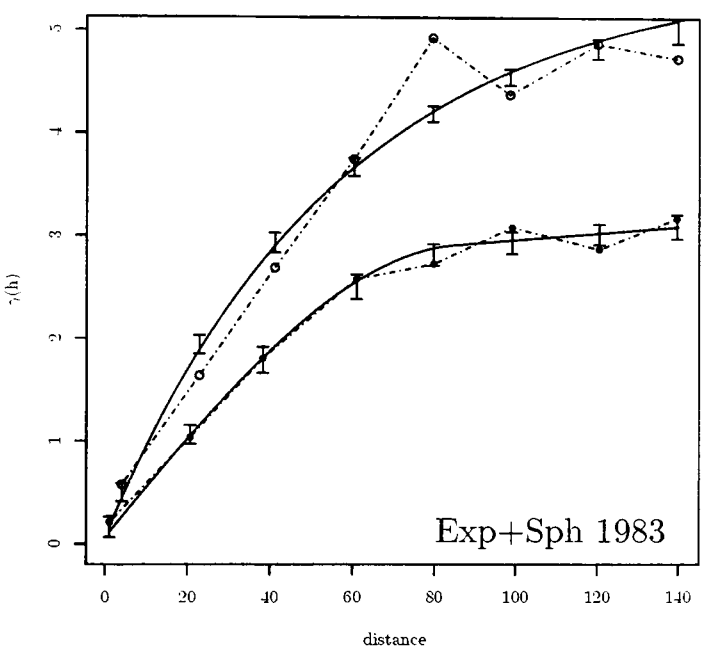
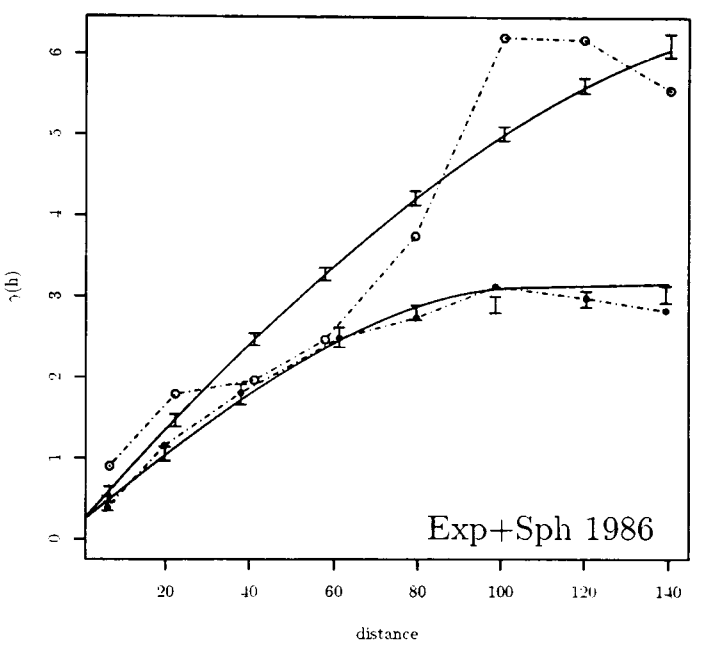


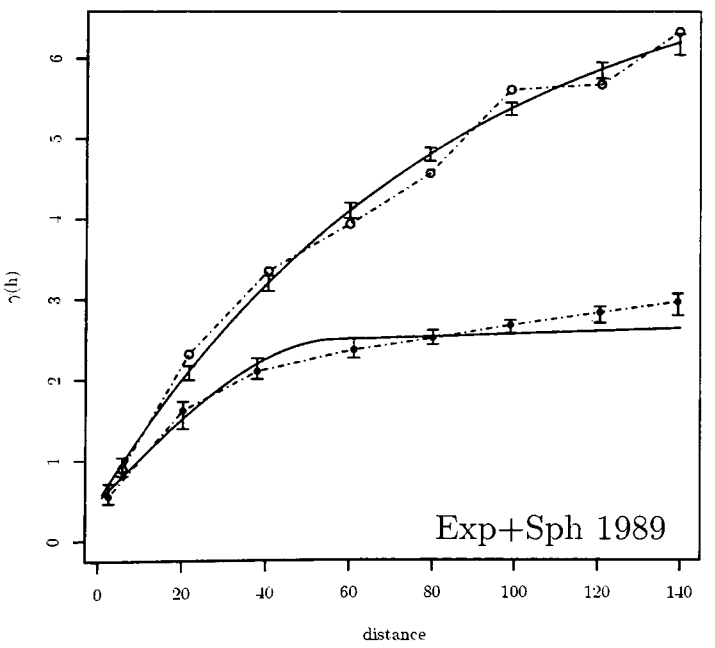
Figure 3.8: Variograms obtained with the spherical model. Plots show the maximum likelihood model and the 95% credibility interval.

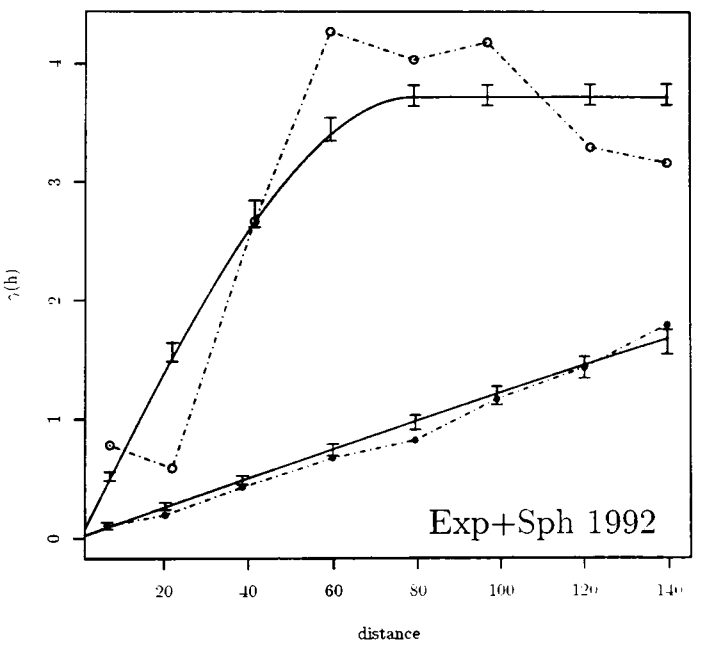

Exp+Sph 1977
γ(h)
distance


Exp+Sph 1980
γ(h)
distance


Exp+Sph 1983
γ(h)
distance


Exp+Sph 1986
γ(h)
distance


Exp+Sph 1989
γ(h)
distance


Exp+Sph 1992
γ(h)
distance

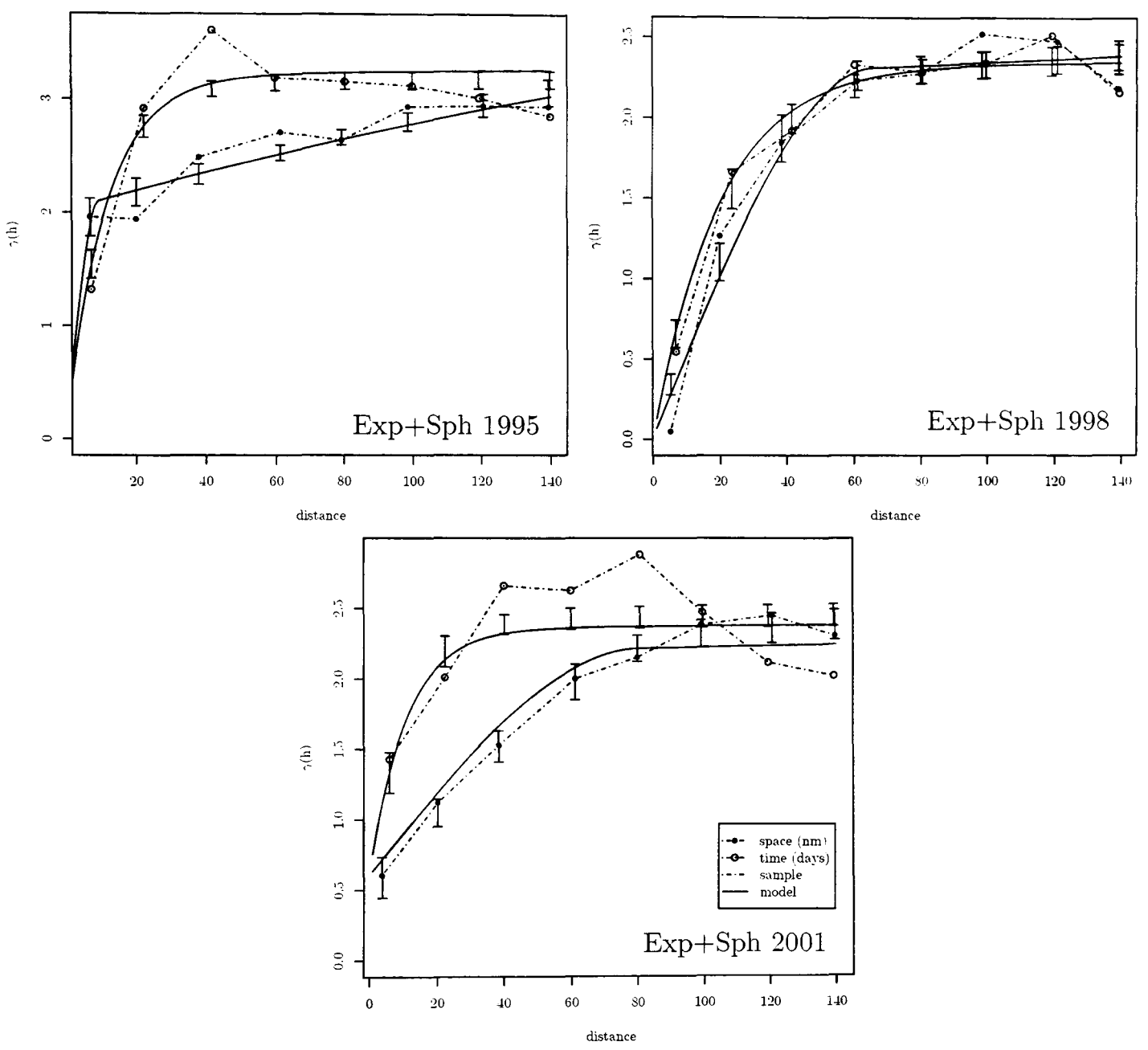


Figure 3.9: Variograms obtained with the nested (exponential + spherical) model. Plots show the maximum likelihood model and the 95% credibility interval.

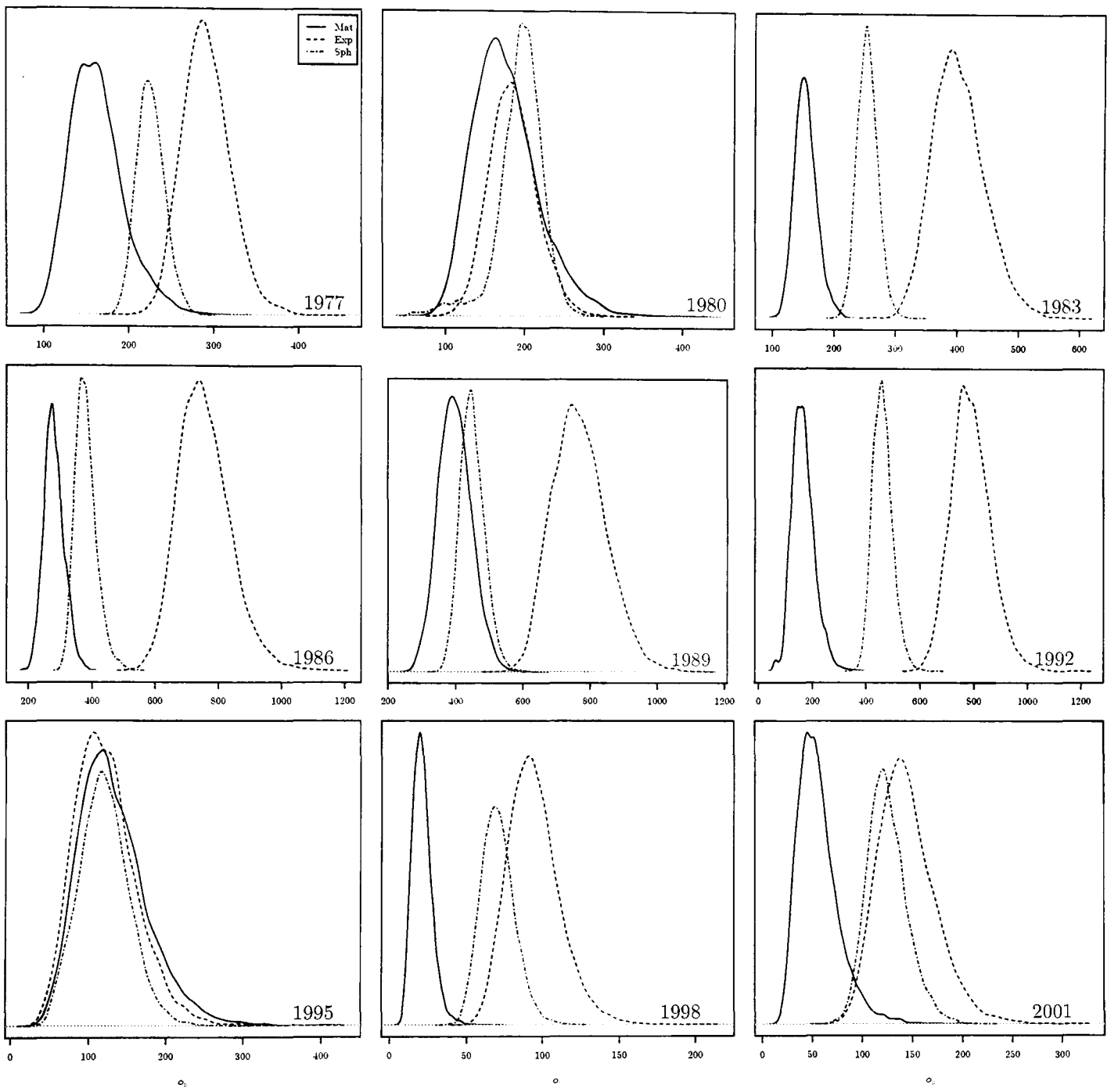


Figure 3.10: Posterior distributions of the spatial range, ϕ_x, for all years and the three non-nested models.

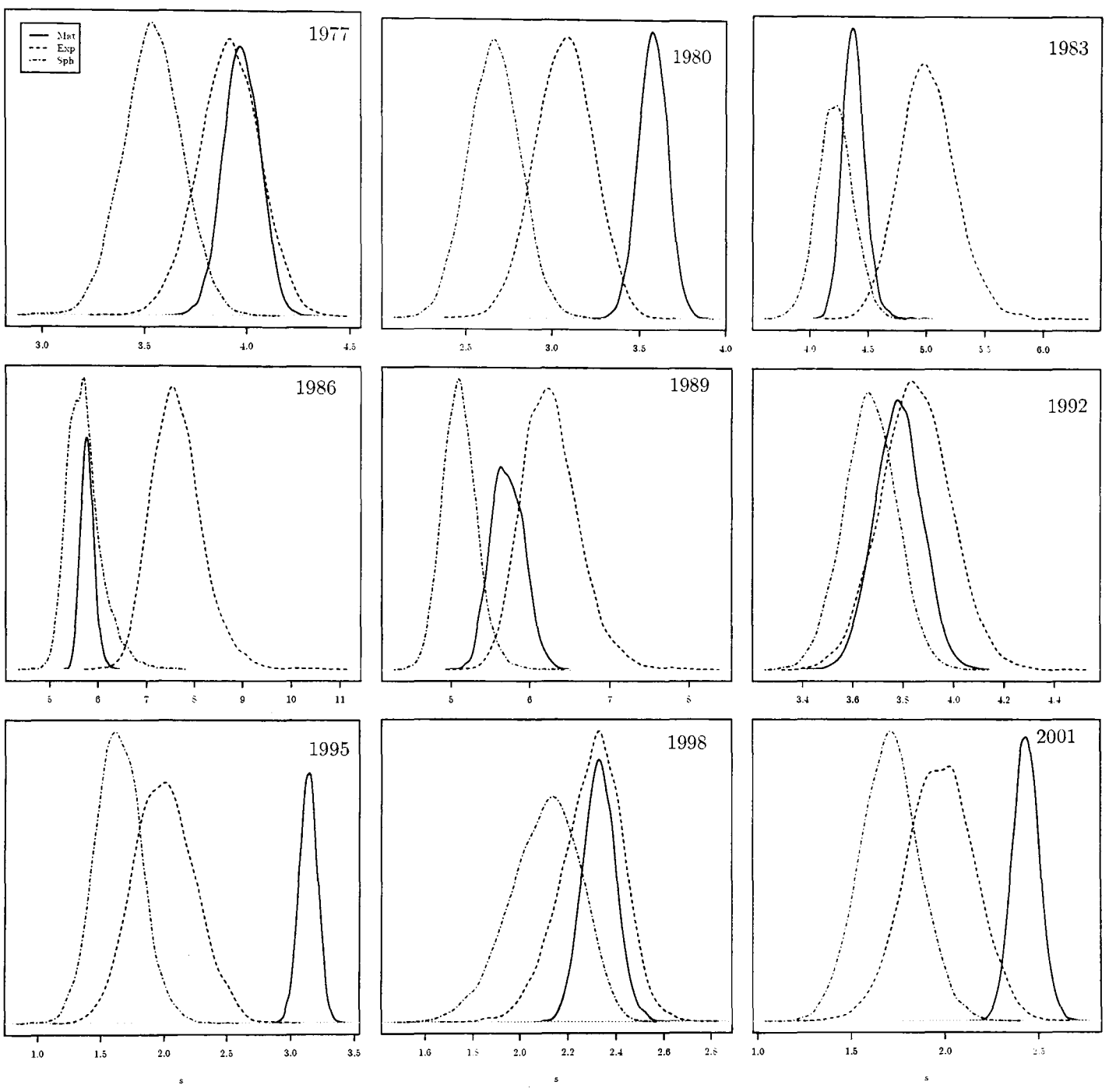


Figure 3.11: Posterior distributions of the sill, C, for all years and the three non-nested models.

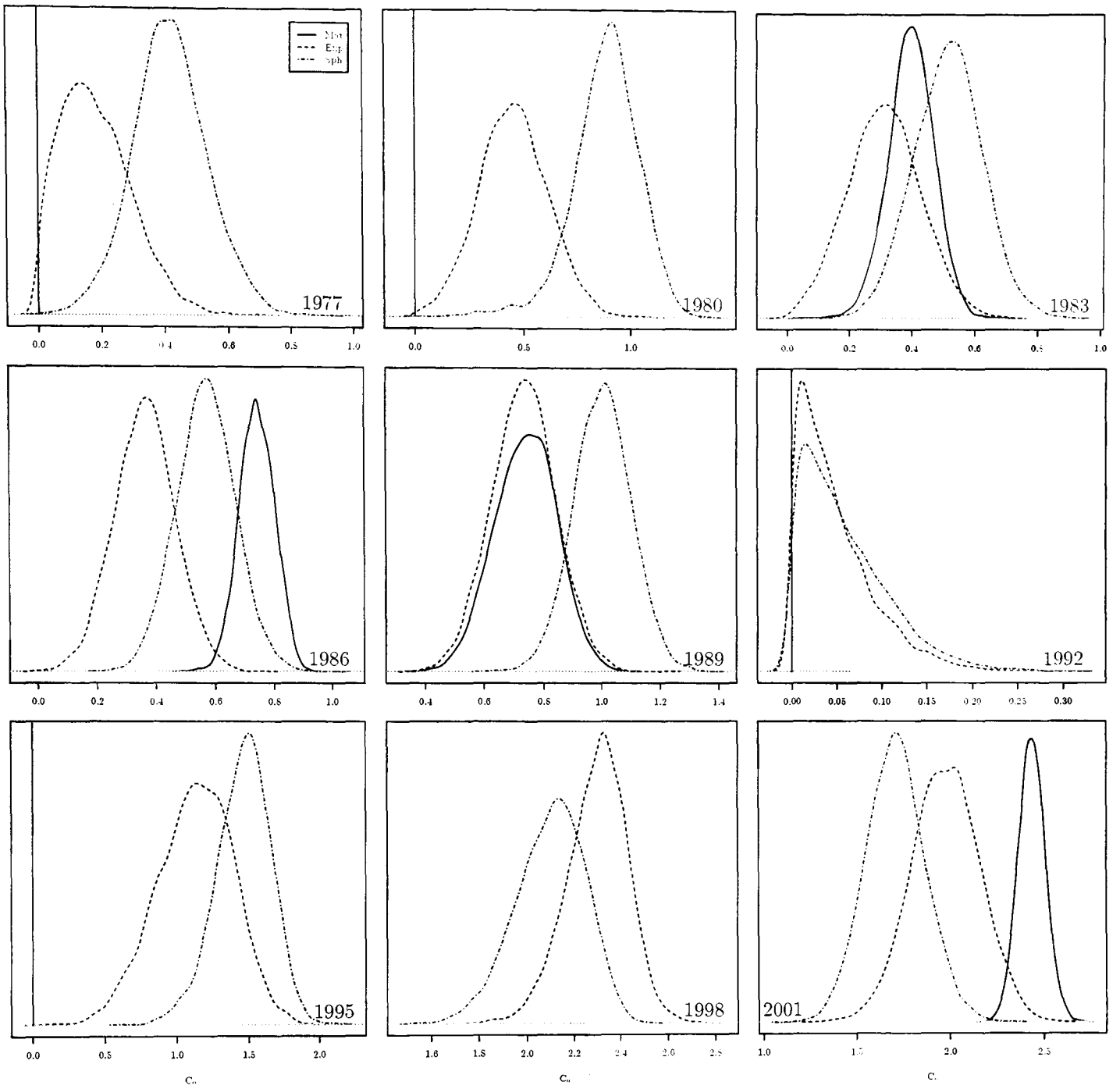


Figure 3.12: Posterior distributions of the nugget, C_0, for all years and the three non-nested models.

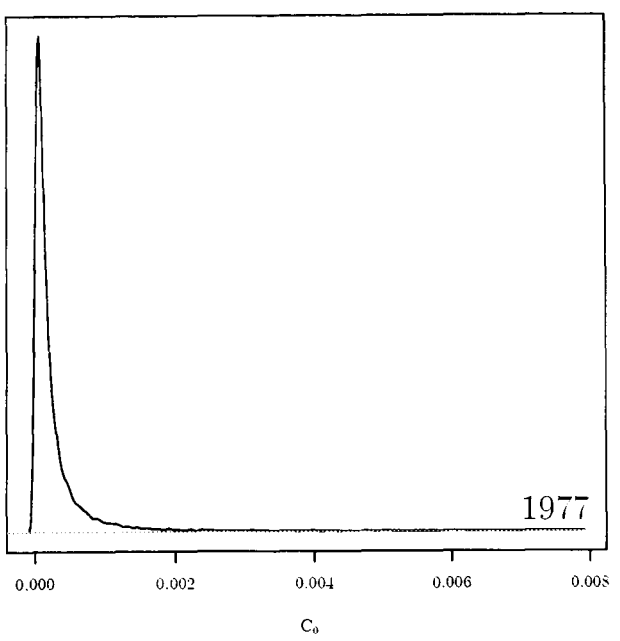


Figure 3.13: Posterior probability for the nugget C_0 of the Matérn model in 1977.

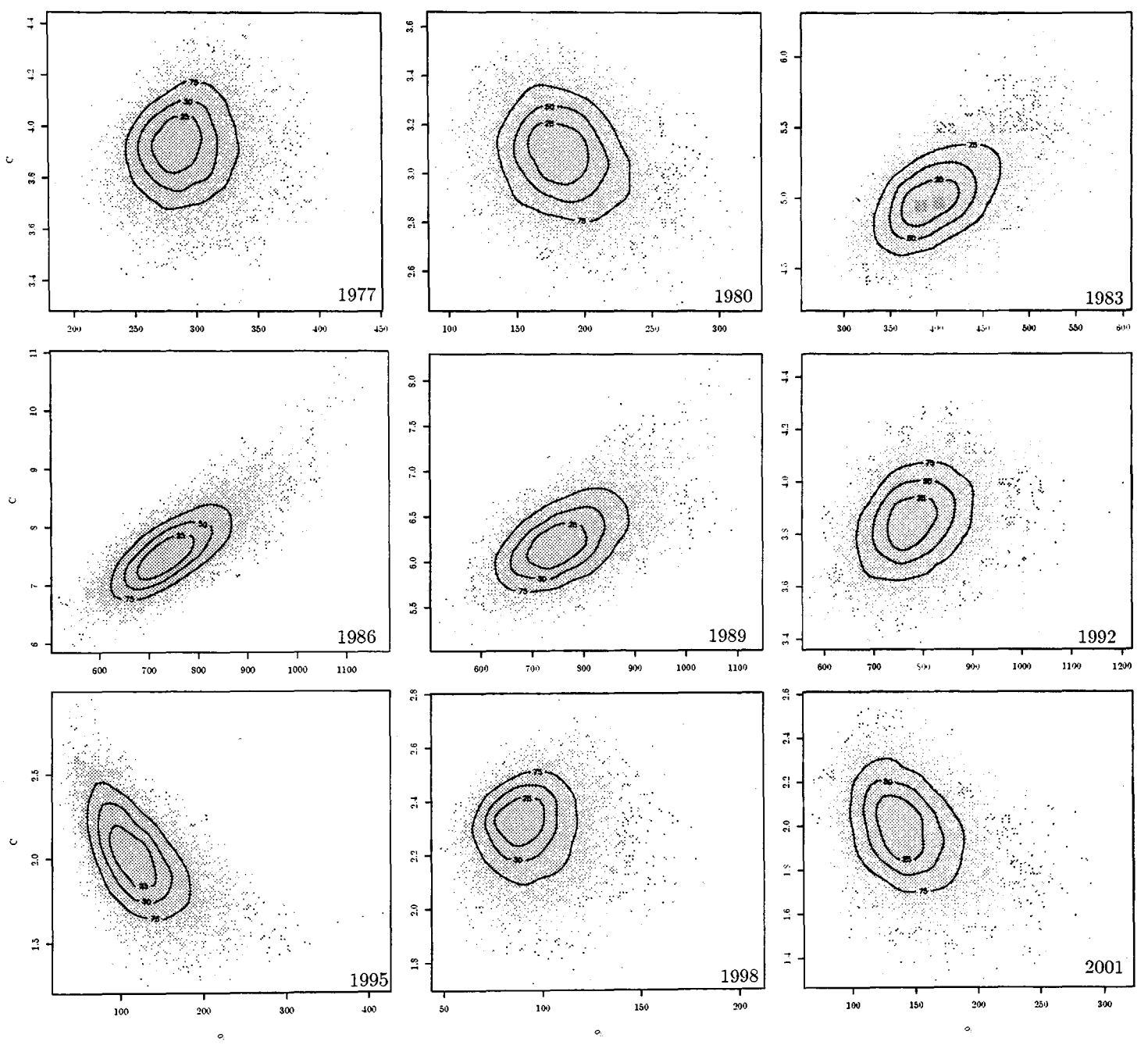


Figure 3.14: Joint posterior distributions of the range in the spatial dimension, ϕ_x, and the sill, C.

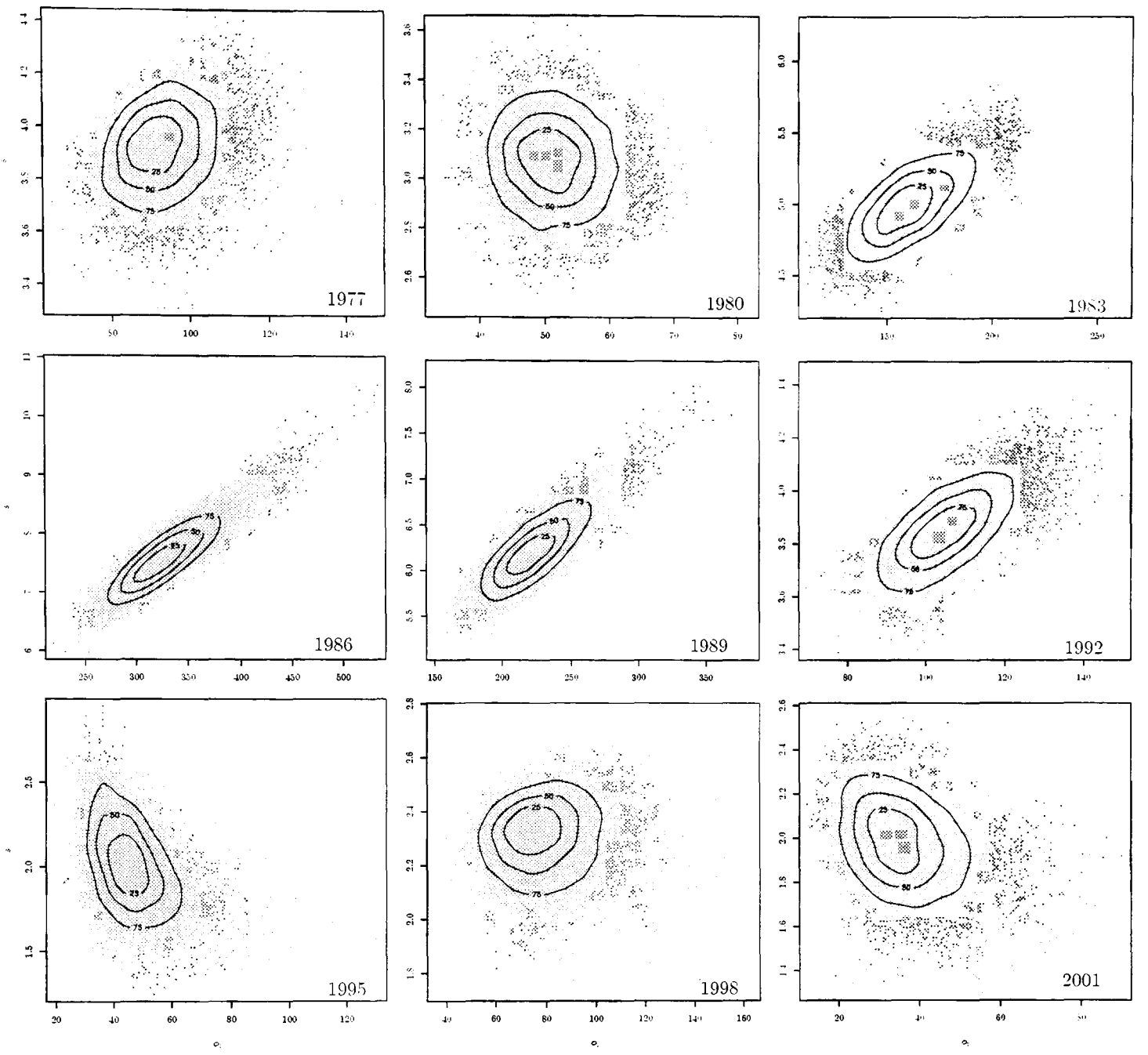


Figure 3.15: Joint posterior distributions of the range in the temporal dimension, ϕ_z, and the sill, C.

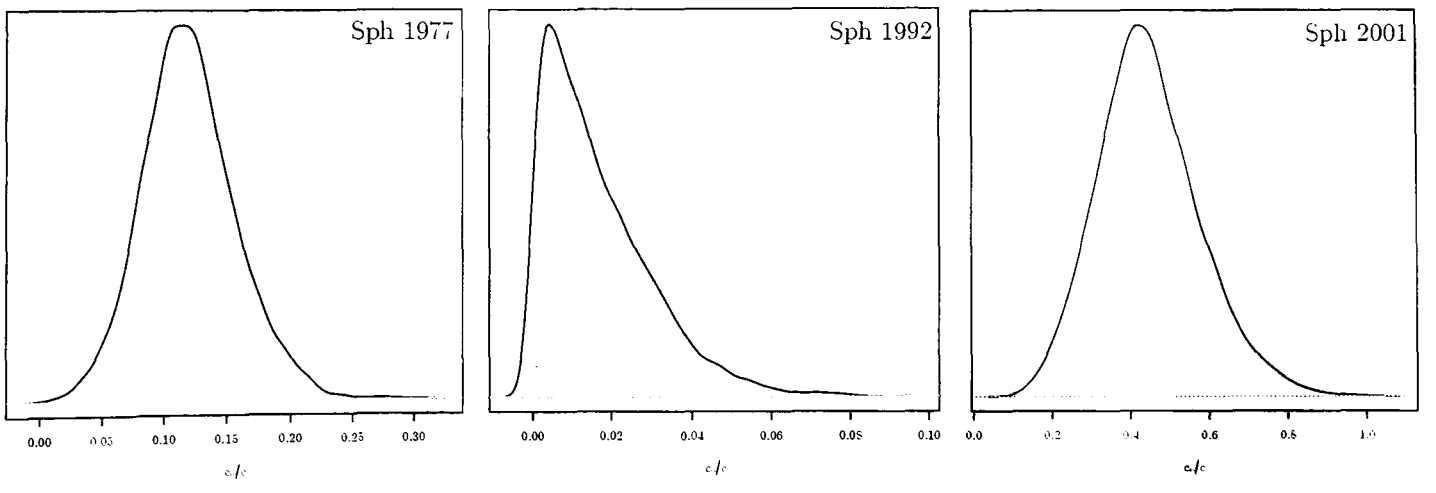


Figure 3.16: Relative nugget posterior probability distributions for the spherical model and in 1977, 1992 and 2001.

3.5 Discussion

The covariance structure of pelagic egg abundance across time and space has been modelled by means of Bayesian semivariogram models. A number of model structures have been fitted to experimental variograms within a Bayesian framework through MCMC methods. The results have been shown to be independent of the choice of prior distributions, and the samplers appear to have arrived at stationary posterior distributions. The hierarchical formulation presented here has emphasised the value of combining different realisations of the same process under a unifying analytic framework.

Although variograms might be considered as merely a necessary step in the use of kriging methods for point or global estimation, they provide by themselves useful information about the spatio-temporal structure of data and about the dynamics of the process responsible for the observed values. In this case, egg abundances are the result of the spawning pattern of a pelagic fish species. This process is influenced by numerous factors, both biotic and abiotic (Freon & Misund, 1999). Schooling or shoaling behaviour, migration routes, and environmental preferences greatly determine spatial distribution. This is subsequently affected by currents, salinity levels and wind patterns. Abundance depends on fecundity and population size, which are in turn affected by environmental conditions prior and during spawning, and natural and fishing mortality. Examination of a variography analysis of multiple surveys, as the one presented here, might present a good opportunity for examining spatio-temporal patterns in spawning under a range of environmental scenarios, and provide significant information on the relative importance of the influencing factors presented above.

Variograms of egg abundance over space and time mostly represent the result of the migratory movement of adult fish during the spawning season. The pattern and dynamics of migratory and schooling behaviours should, for example, have a direct influence in the shape and range of the variograms. In a comparison between anchovy and sardine eggs sampled along the Californian coast, Curtis (2004) noted the apparent relationship between the superior swimming ability of sardine, which tends for form denser aggregations, and the observed variograms with large spatial autocorrelation and small nuggets. In contrast, anchovy egg variograms showed larger nugget values and autocorrelation levelling off at shorter distances. Another important factor that can also influence variogram shape is population abundance and its density-dependent effect on fecundity (Barange *et al.*, 1999). More fecund species can be expected to expand their range as population size increases, instead of simply increasing their density as tends to be the case if individual fecundity is lower.

Atlantic mackerel is a highly fecund and extremely mobile species: a 275 g female can produce approximately between 375,000 and 500,000 eggs per season (ICES, 2004b), and NEA mackerel are known to migrate between the Basque Country and Norway during a single spawning season (Uriarte & Lucio, 2001). Changes in population size may therefore be expected to affect its spawning area more significantly than its density.

Female condition before spawning has been found to vary significantly among years (Slotte, 2004), and this might have a direct effect on potential fecundity in the following spawning season. Condition seems to be dependent on *Calanus* abundance, probably driven by environmental and oceanographic factors. But population size could also be an important factor here, as wintering in dense aggregations could make NEA mackerel more susceptible to density-dependence effects on fecundity. Although a direct link between winter condition and potential fecundity has not been established (ICES, 2004b), the complexity of the matter appears to void simple interpretations of the observed changes in variogram shapes. Furthermore, possible adaptations of life history to environmental changes have been proposed for other small pelagics (MacCall, 2001). If proved, they would complicate even further the interpretation and prediction of changes in spawning.

One of the results obtained that can be directly interpreted is the relative nugget. Sampling variance and small-scale variability both contribute to increase the variogram nugget. Relative nugget values across all years, as inferred from the mean nugget hyperparameter, represents between 2 and 10% of the variance. This low value might be considered as a sign of reasonable accuracy in the present sampling protocol. But the estimation of actual sampling error is hampered by the lack of replicates samples taken over a short time around the same area. An attempt to invetigate this was carried out, albeit in a very tentative manner, during the 1998 survey (C. Hammer, pers. com.). A total of 10 samples were taken within the same sampling square in a period of 20 hours. The coefficient of variation reached 42%, a value significantly higher than the one inferred here from the variograms. Consequently, the relative nugget in this case might be taken as a very conservative estimate, probably unrepresentative of the real degree of variation present in the sampling procedure. A possible interpretation for this result is that with the sampling design that generated these data, the estimated nugget might be better taken as an estimate of the variation likely to be observed between neighbouring grid squares.

In contrast with other published studies, this analysis covers a relatively long series of surveys, separated by nearly 15 years. Many uses of geostatistics in fisheries concentrate on the analysis of a single survey (Fletcher & Sumner, 1999, for example). Changes in sampling methods, intensity and coverage are likely to have an important part in the observed variation in variogram shapes and parameter values among years. These are larger than those found by some authors for different species with divergent spawning ecology (Curtis, 2004). Correlations between variogram parameter values and a number of oceanographic processes that could be responsible for the observed variation has not proved conclusive (C. Imrie, pers. com.). We might then have to conclude that we are observing here the result of a process, mackerel spawning, subjected to multiple influences in its exact development. The combination of environmental, biological and human factors could all be influencing the precise dynamics of spawning in very complex ways. But despite the manifest differences in which mackerel spawning occurs every year, the general patterns are still consistent from a biological perspective.

A number of model selection criteria have been tested on the four proposed variogram models. More parsimonious models, those with fewer parameters, appear to fare better. The Matérn model, with all its potential in a hierarchical setting, gives rise to surprising variogram models for some years. Model selection based on a single, or a limited number of performance mesaures is not always possible or advisable. In the case of hierarchical models that do not form a nested family, this comparison is further hampered by the difficulty in estimating the effective number of parameters, as discussed before. Nevertheless, the results presented above appear to allow selecting, from a practical point of view, one of the four models tested. The nested exponential and spherical model provides good fits for the whole range of variogram shapes encountered, in both the time and space dimension. Despite the increased number of parameters, its model fit metrics are some of the best (Table 3.3). Behaviour at the origin is also an important factor given the fine grid that will be used for kriging, as presented in Chapter 4. The nested model does not suffer of the problems encountered with the Matérn model. User of the later at short distances, with its flat shape, would greatly increase the screen effect and difficult kriging estimation. The nested model was thus choosen as the best representation of the experimental variograms obtained and the most promising for kriging estimation of egg production. Kriging was nevertheless carried out for all models, in order to uncover the relationship between variogram model characteristics and kriging estimates for this dataset.

Variograms, in common with other models, might be constructed with two different objectives: explanation or prediction. In the first case the intention is to provide a coherent explanation of reality. We have attempted in this chapter to show how variograms can be used to help us develop and test various explanations of natural phenomena. Hypotheses on the likely patterns of pelagic fish migration and spawning can be explored through variogram modelling. For the second objective, prediction, the approach is to select the best possible model to direct our prediction procedure, kriging in this case. But variogram model selection based on the fit to the sample variogram only does not consider its ultimate purpose. The question that needs answering is whether the deficiencies in the model have any effect on its substantive inferences (Gelman *et al.*, 1995). In this case, the apparent advantages of certain variogram models presented before cannot be completely clarified until tested for spatio-temporal estimation through kriging.

Chapter 4

Monte Carlo block kriging of egg abundance

Estimation of variable values at unsampled locations is accomplished in geostatistics with the help of a model for spatial, and sometimes temporal, dependence. A variety of such models were developed in the previous chapter, and will be applied here to the problem of estimating the value of our variable of interest, egg production, over the whole space-time volume. Egg production is estimated at all points within a regular grid based on the aforementioned model. Estimates for restricted sections of the space-time volume are then combined to arrive at a final estimate of Annual Egg Production. The procedure of block kriging (Cressie, 1991) will be applied here to this task.

The spatio-temporal covariance models developed in this study incorporate the uncertainty in the estimation of their parameter values, and this will be carried on to the next step in the estimation process, the kriging procedure. Kriging will be carried out in a Monte Carlo simulation setting, in which parameter values will be drawn from the joint posterior distribution of the variogram parameter set, and kriging carried out for each of those sets of parameter values.

The results of this exercise will be presented first in terms of the variability in AEP estimates for NEA mackerel. A comparison between the results obtained under each alternative model will also be set forth. Further comparisons will be made by use of simulated datasets. Finally, results of the Monte Carlo block kriging procedure will be used to further analyse and understand spatio-temporal patterns of egg spawning in NEA mackerel.

4.1 Three-dimensional block kriging

Kriging is the name adopted by geostatisticians for a family of generalised regression algorithms originally developed for solving problems of ore reserve estimation (Goovaerts, 1997). The term was coined by one of the initial developers of the field, G. Matheron, in recognition of the pioneering work of a South African mining engineer, Danie Krige. The problem of estimating the value of an attribute z at any unsampled location u by using

the available data on z over the study area, is approached here by the use of a linear regression estimator $Z^*(u)$, defined as

$$Z^*(u) = \sum_{a=1}^{n(u)} \lambda_a(u) Z(u_a) \tag{4.1}$$

where $\lambda_a(u)$ is the variogram-determined weight assigned to datum $z(u_a)$. interpreted in this case as a realisation of the random variable Z at location u_a. These weights are obtained by applying the previously fitted variogram model (Chapter 3) to the distances between the gridded points and the available datapoints. The number of points involved in the estimation is represented by $n(u)$, and is commonly limited to those found within a given neighbourhood or window, $W(u)$.

4.1.1 Block kriging

When interest lies in the estimation of the value of a certain variable over an area or volume, a variation on ordinary kriging, called *block kriging*, is commonly applied. Block kriging is a generic denomination for the estimation of average values over a segment, surface or volume, V, in contrast with *point kriging*, which refers to estimation of point support (Goovaerts, 1997). The average value of a variable z in block V centred at location u, can be estimated as

$$z_V(u) = \frac{1}{|V|} \int_{V(u)} z(u')du' \approx \frac{1}{N} \sum_{i=1}^{N} z(u'_i) \tag{4.2}$$

where the second section of the equation shows an approximation to the integral, as a discrete sum of values defined at a number of points u'_i within block V. The block in Figure 4.1.1 is, for example, discretised by four points, u'_1 to u'_4. The value of z at block V, $z_V(u)$, can then be estimated as the linear average of the kriging estimates at the discretising locations (Carr & Palmer, 1993). Values at each of the four points in this block are in turn estimated following equation 4.1 from the combination of a variogram model and available datapoints.

Applying block kriging to the problem of global estimation is carried out in two steps. First, an average value is estimated for each defined block, following the method outlined above. Second, the global estimate is calculated as a linear combination of those block estimates, each weighted proportionally to the area or volume of the block (Isaaks & Srivastiva, 1989; Cressie, 1991). Back-transformation is then carried out on the log-transformed results to obtain an estimate of annual egg production.

4.1.2 Spatio-temporal kriging

Kriging has been often applied to the estimation and analysis of static phenomena, such as ore deposits. Spatial interpolation in two or three dimensions was thus sufficient in those

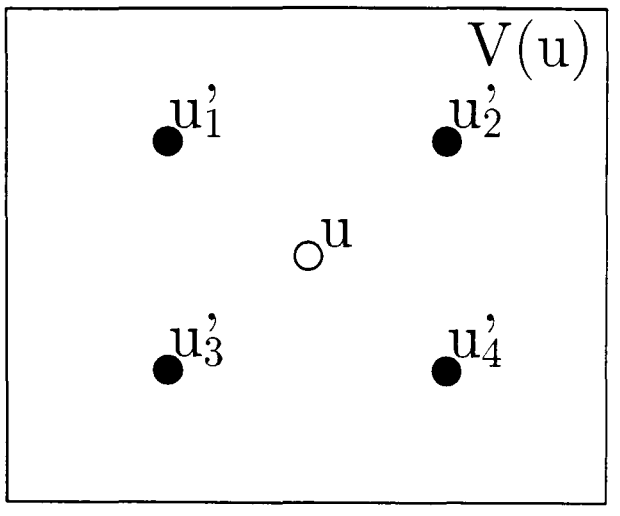


Figure 4.1: Discretisation procedure in block kriging. A two dimensional block V centred on location u, and four discretisation points, u'_1 to u'_4.

situations. Other fields are concerned with the analysis of a dynamic process, requiring the temporal dimension to be explicitly considered. Such is the case in most fisheries applications, including egg surveys. In the "traditional" AEP estimator as applied for ICES mackerel surveys (ICES, 1999a), temporal interpolation is achieved by considering a number of point estimates of instantaneous egg production along the year and integrating below the curve generated. But in order to arrive at those point estimates, egg abundances collected along a single sampling period, sometimes covering a period as long as four or five weeks, must lose its temporal character and be considered to provide a snapshot of egg production in the middle day of the period (See Chapter 2 for a more complete explanation).

In contrast, modelling techniques that explicitly consider the actual time component of each datapoint should have the ability to provide more accurate estimates of egg production, allow detailed investigation of temporal dynamics in spawning, and separate the confounding effects of space and time on egg abundance (Augustin *et al.*, 1998). Geostatistical methods can be applied to the problem of modelling spatio-temporal phenomena, although this is at present an area of continuous development (Todini, 2001). Two main approaches have been identified for developing geostatistical space-time models, the first by considering a single spatio-temporal random function, and the second involving vectors of spatial random functions or of time series. An introductory review of available models for either approach can be found in Kiridiakis & Journel (1999).

Some authors have applied state-space formulations of the linear kriging estimator to incorporate the temporal component to problems of spatial interpolation (Stroud *et al.*, 2001; Huerta *et al.*, 2004). This approach appears to be best suited to problems such as those encountered in meteorology, where sampling takes place almost continuously in a number of fixed stations. Temporal variation in each cell can be modelled by means of time series methods. Spatial interpolation is in turn concerned with changes in the temporal dynamics observed at different locations. These methods also require data on

the value of a related variable at all locations, commonly one that is easily recorded such as temperature. The model is thus dependent on the availability and the strength of the relationship between the variable of interest and the chosen covariate.

Following the space-time separable approach applied in the present study to the problem of variogram estimation, kriging of pelagic egg abundance has been considered here equivalent to a three-dimensional spatial situation, where time takes the place of depth or height. It is important to note here that the procedure presented does not consider the space-time covariance structure as providing a complete explanation of the spatial and temporal dynamics of fish spawning, but that merely allows determination of the relationship between egg abundances at various points in space and time. In addition, although spawning might appear differently when observed from the temporal or spatial perspectives, the patterns observed are in fact the result of a single process, namely fish migration during the spawning season.

4.1.3 Implementation of the kriging procedure

A spatial grid extending beyond the limits of the known distribution of mackerel spawning was set up. Block or cell dimensions were set to 15×15 nautical miles and 7 days. Figure 4.2 shows the spatial grid, the centre points of each block, on a location map. The block kriging procedure estimated egg abundance values at 6 points on each dimension, making a total of 216 points per block.

Kriging was carried out using the *gstat* software (Pebesma & Wesseling, 1998). Code was written to sequentially generate kriging and cross-validation input files from the MCMC output files, run *gstat* on them, and process the output, for each simulation, year and model.

4.1.4 Search neighbourhood

Common practice in kriging is that the search for data points to use when estimating the variable value at an unsampled point is limited to those closest to the location being estimated. The search neighbourhood, $W(u)$, can be selected both in terms of area, or volume, around the point, or as a maximum of data values to use, starting with those in closer proximity (Cressie, 1991). In addition, the minimum and maximum number of data values to include in the estimation at each point must be specified. The reasons for this restriction include considerations of computational efficiency (Goovaerts, 1997). For example, the size of the kriging system increases following approximately a cubic relationship with the number of points to estimate (Pebesma & Wesseling, 1998).

The search neighbourhood, defined either as search distance or as number of points, was not incorporated in the present model as an estimable parameter, not was the uncertainty possibly introduced by its choice formally quantified. Tests to evaluate the impact of the values chosen were carried out on a single dataset using different search radii and/or minimum and maximum number of points. The kriging procedure appeared from those tests to be more sensitive to the choice of search radius in the temporal dimension. High

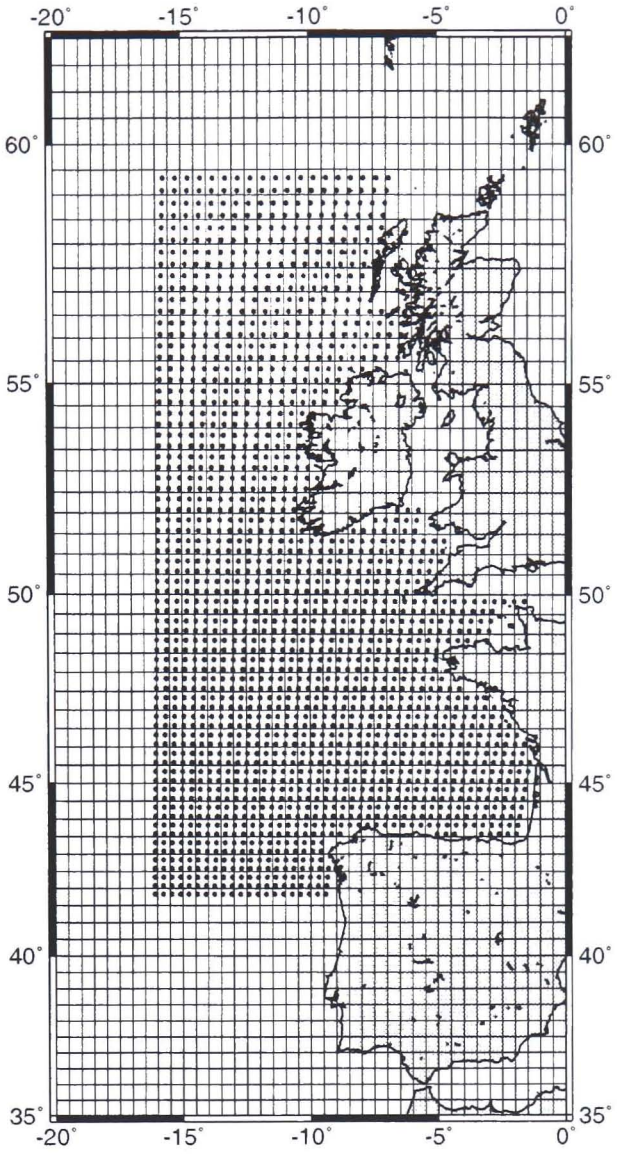


Figure 4.2: Center points of the 15x15 nautical mile blocks defined for kriging.

values, if found in areas that were not sampled again in previous or subsequent weeks, apparently exerted their influence for a longer time period than could be reasonably expected from the spawning biology of mackerel. For example, Figure 4.3 shows the difference in weekly abundances estimated for the same variogram model with two values of search radius in the temporal dimension (100 days and 50 days). Estimates of AEP for those two values for the search radius were almost identical ($1.2356 \cdot 10^{14}$ vs. $1.2355 \cdot 10^{14}$), as the variograms already give data points found beyond 40 or 50 days a very small weight (See the variogram models in Figures 3.6-3.9). However, the search radius in the temporal dimension was nevertheless reduced from an initial value of 100 days to 60 days, considered to be more in accordance with development stages of mackerel eggs and the migratory behaviour of Atlantic mackerel (Coombs *et al.*, 2001: Uriarte & Lucio, 2001). The values finally used are summarised in Table 4.1.

The search algorithm implemented in the *gstat* software package does not allow the specification of separate values of the search neighbourhood parameters in each direction. Values along the z axis were thus re-scaled from days, d, to an arbitrary unit of time $t = 1.6d$. Results were then scaled back to days and weeks for presentation.

Table 4.1: Parameter values used for the searching procedure in both dimensions.

Plane	Minimum number of values	Maximum number of values	Search radius
Spatial (x,y)	3	20	100 nm
Temporal (z)	3	20	60 days

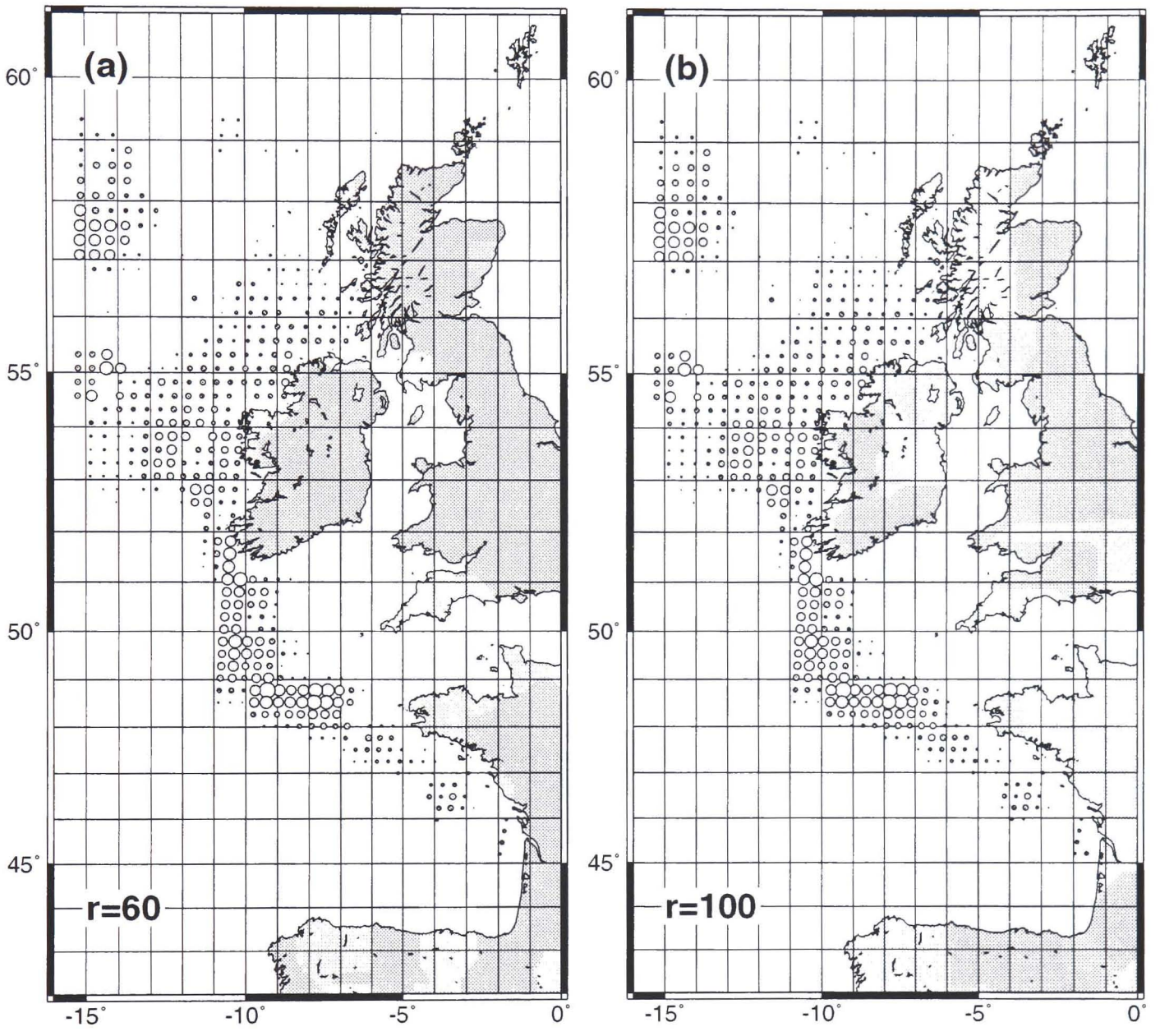


Figure 4.3: Maps of weekly egg production estimated by block kriging using different temporal search radii: 60 days (a), and 100 days (b).

4.2 Monte Carlo kriging estimation

Considering the whole kriging system in a Bayesian setting is not without merit, and various studies (Omre, 1987; Handcock & Stein, 1993) have presented Bayesian formulations of the kriging estimators that include the variability at both steps: variogram modelling and kriging. But given the characteristics of the kriging procedure necessary for the estimation of pelagic egg production, a simplified approach was developed here. A Monte Carlo procedure was implemented that drew on the Markov chains obtained from the Bayesian hierarchical variogram models. For each model and year, 500 iterations of the

final section of the Markov Chains were selected at distances of 100 iterations. This value was chosen after inspection of the autocorrelation plots for each chain, and is larger than the longest lag with autocorrelation greater than 0.3 in the worst-mixing chain. Block kriging was then conducted using the variogram model parameter values from each of those draws and a distribution of egg production values, both at each grid location and for the whole area-time volume, was obtained. Draws for the MCMC output were carried out without consideration of the individual posterior probability obtained by each of those parameter sets. The assumption here is that the end section of the MCMC chain is perfectly mixed, so drawing from it yields values whose probabilities are in accordance with those observed in the variogram models.

4.2.1 Implementation details

The Monte Carlo procedure outlined above consisted of a total of 500(MC iterations) $\times$ 9(years) $\times$ 4(models) = 18,000 iterations. Each run required between 2 and 5 minutes of computing time on a fast personal computer. A decision was taken to implement the procedure to run in a cluster platform consisting of 50 interconnected processors, belonging to the Centro de Supercomputación de Galicia (CESGA), based in Santiago de Compostela. Code was written specifically to run on this platform, that automatically distributed the simulations among the available processors and stored the results from each run.

The complete output generated by the Monte Carlo kriging exercise quickly grew to a size that complicated its analysis. An array able to hold the full set of results would consist of six dimensions (longitude, latitude, week, iteration, year, model), and would need to have a total size of $72 \times 40 \times 28 \times 500 \times 9 \times 4 = 1{,}451{,}520{,}000$ or nearly a billion and a half cells. Such an structure quickly exhausts the memory and processing capacity of most personal computers. Analyses were thus restricted to sizeable sections of the array, a fact that sometimes limited what results and summaries could be presented. A more efficient storage mechanism could possibly have been achieved by the use of an external database server to which queries could be directed by the analysis procedures. The relative complexity of designing such a database, and the limited number of statistical operations that most relational database servers offer, precluded the use of this mechanism during this study. Future studies along similar lines to this one should carefully consider efficient storage and access of simulation results.

Cross-validation was carried out to obtain a correction factor for the log-transformed egg abundances. Kriging of log-transformed values is known to bias the final results, due to the lack of robustness of the logarithm with regard to the multivariate normal hypothesis inherent in Gaussian kriging (Journel & Huijbregts, 1978). A corrective factor, K_0, can be obtained by equating the mean estimated values at those locations where data are available to their known values through cross-validation kriging (Goovaerts, 1997). Each datapoint is temporarily removed and its value estimated by kriging using the remaining data. The same procedure is then carried out for all datapoints. The corrective factor, K_0,

is estimated as the mean ratio between sampled and estimated value. In this case the full distribution of deviations were also stored for subsequent investigation of the uncertainty introduced by this step in the estimation procedure.

Results from the *gstat* software, consisting of egg abundances and kriging variances at all grid locations, were then rescaled back to days and weeks. Log-abundances were back-transformed and the abundance estimate for each point in the spatial grid was then multiplied by the area of the grid square. The final values were obtained by multiplying the back-transformed egg abundances by the mean correction factor.

4.3 Kriging estimates of North East Atlantic mackerel Annual Egg Production

The estimates of AEP obtained for every year and variogram model are summarised in Table 4.2. The figures shown are the median of the 500 kriging estimates, together with their corresponding coefficients of variation, estimated as cv $= \sqrt{e^{\sigma^2} - 1}$. Variability appears to be higher for the Matern and spherical models, with average cv of 42% for Matern, 38% for spherical, 32% for nested and 31% for exponential. The differences, however, are in concordance with the variability observed in the variogram models themselves.

Table 4.2: Estimates of NEA Atlantic mackerel annual egg production by year, in terms of median AEP$\times 10^{15}$ and CV, for all models in this study. Models are Matern (Mat), exponential (Exp), spherical (Sph) and nested (Nes). Nested refers to the Exponential + Spherical model.

	Mat		Exp		Sph		Nes	
	median	cv	median	cv	median	cv	median	cv
1977	1.82	0.39	1.89	0.3	1.67	0.44	1.48	0.45
1980	1.93	0.43	1.88	0.47	1.83	0.33	1.76	0.39
1983	1.79	0.52	1.56	0.49	1.51	0.35	1.61	0.41
1986	2.04	0.68	1.64	0.29	1.68	0.44	1.45	0.25
1989	2.23	0.33	1.66	0.27	1.65	0.46	1.63	0.37
1992	1.89	0.41	2.01	0.32	1.88	0.41	1.99	0.27
1995	1.65	0.36	1.57	0.23	1.54	0.38	1.68	0.32
1998	1.4	0.32	1.47	0.25	1.36	0.35	1.48	0.22
2001	1.23	0.28	1.25	0.19	1.09	0.25	1.36	0.23

Comparison of the results obtained in the present study and those generated by ICES Working Group on Mackerel and Horse Mackerel Egg Surveys are presented in Table 4.3. Discrepancies are notable for some years, for example 1986 and 1989. They appear to be higher at the start of the survey series where coverage could be considered to be relatively sparse, as sampling attempted to cover larger areas without an apparent increase in resources. In those situations, a geostatistical estimator is likely to interpolate over a larger number of spatio-temporal cells, thus giving a higher estimate of egg abundance, as higher values influence kriging over larger areas.

Table 4.3: Comparison of kriging estimates of annual egg production by year with those obtained using the *traditional* estimator. Results show the median estimates from kriging for each variogram model in eggs$\times 10^{15}$, and the percentage difference from the Working Group estimate, as presented in ICES (2004b). The last line presents the mean absolute deviation for each model. Years marked with * were modelled in a separate variogram model.

	Mat		Exp		Sph		Nes		**Trad**
	median	diff	median	diff	median	diff	median	diff	
1977	1.82	-8.1	1.89	-4.5	1.67	-15.7	1.48	-25.3	1.98
1980	1.93	4.9	1.88	2.2	1.83	-0.5	1.76	-4.3	1.84
1983*	1.79	17.0	1.56	2.0	1.51	-1.3	1.61	5.2	1.53
1986*	2.04	64.5	1.64	32.3	1.68	35.5	1.45	16.9	1.24
1989*	2.23	46.7	1.66	9.2	1.65	8.6	1.63	7.2	1.52
1992	1.89	-2.6	2.01	3.6	1.88	-3.1	1.99	2.6	1.94
1995	1.65	10.7	1.57	5.4	1.54	3.4	1.68	12.8	1.49
1998	1.4	2.0	1.47	7.3	1.36	-0.7	1.48	8.0	1.37
2001	1.23	1.7	1.25	3.3	1.09	-9.9	1.36	12.4	1.21
mean		17.58		7.75		8.74		10.53	

The four variogram models tested provide different estimates of the uncertainty around AEP. Figure 4.4 presents the median and 95% credibility intervals obtained for the four variogram models. At this scale median estimates do not differ greatly, when compared with the differences observed among the estimates of variability. The upper quantile for the Matern model reflects the greater uncertainty already observed in comparisons of the four variogram models (Table 3.2).

The full probability distributions of the kriging estimates of AEP obtained for each year and model are presented in Figure 4.5. The differences among models are remarkably small, with the possible exception of 1986 and 1989. On the first one of those, the Matern model appears to be estimating lower egg production, albeit producing at the same time wider confidence intervals. This same model deviates from the other three in the estimate of the 1989 AEP, although in the opposite direction. The variograms obtained for these two years were considerably linear, and even more so near the origin. Fit to the temporal variogram for 1986 was especially poor. This appears to be having an impact in the kriging estimates for the Matern model, more susceptible to the screen effect, which in turn could become more pronounced in years where a few large egg production values exert an important influence in the final AEP estimate.

4.4 Impact of model choice

The differences among the four variogram models observed in the previous chapter are likely to affect the kriging estimates of egg production that can be obtained from each from them. Annual egg production for each of the four models show remarkable differences for some years (Table 4.2). To further investigate those differences, estimates can be

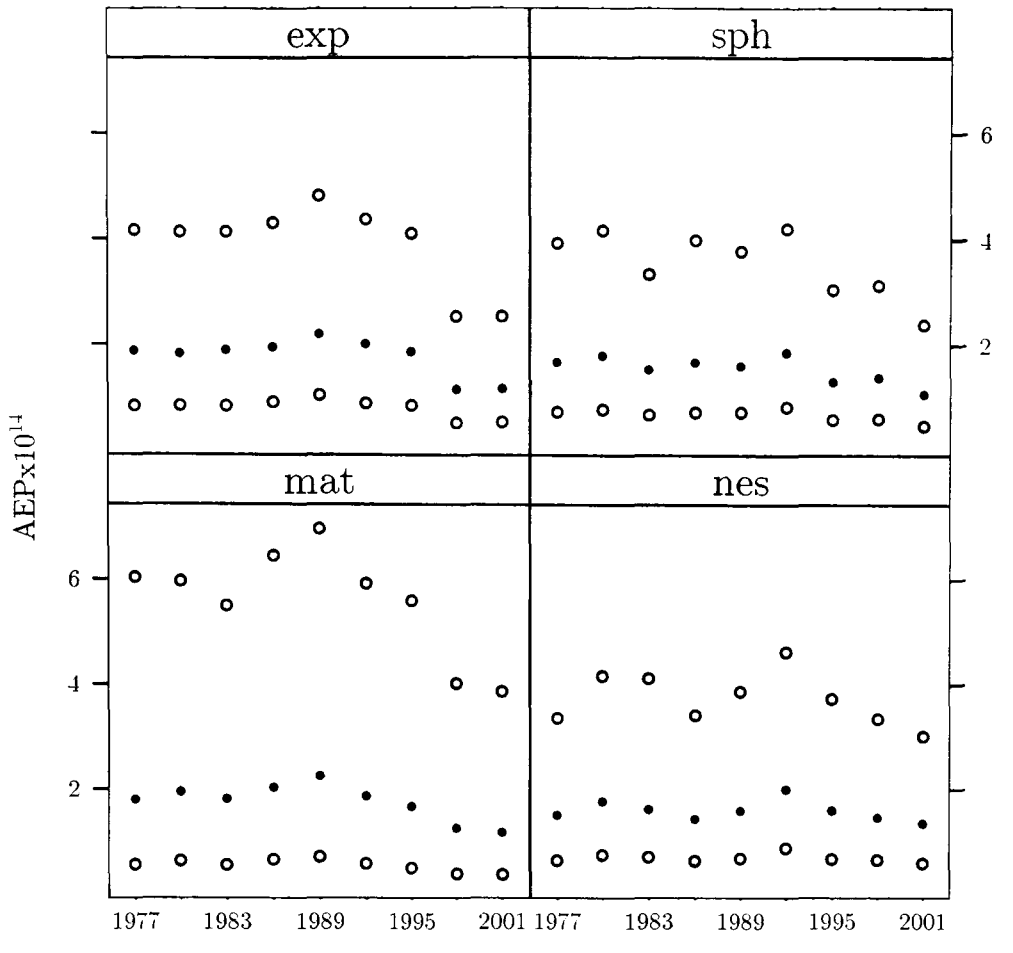


Figure 4.4: Estimates of median annual egg production (AEP), black dots, with their corresponding 95% credibility intervals, open circles, for each variogram model: exponential, spherical, Matern and nested (exponential + spherical).

obtained at a finer temporal level, weekly in this case, and again for all models. Figures 4.6 and 4.7 show the weekly estimated egg production, together with the 95% credibility intervals, for two years with different variogram shapes, 1986 and 1998 respectively. These results further corroborate the model preference outlined in the discussion of the variogram modelling. The patterns in weekly production obtained from the Matérn model are not in agreement with the observed temporal evolution of raw egg abudance, and the uncertainty uncovered appears to be overestimated for certain periods. Most likely this is the effect of the parabolic shape of the model at short ranges. Again the nested model appears to better capture the known temporal dynamics of spawning: sharp increase of activity at the start of the season, peak at or before mid-season, and slow decline of production.

4.4.1 Generation of simulated datasets

The use of simulated datasets is a common procedure for testing statistical estimators of random variables in a variety of contexts (Restrepo, 1995; Bogaert, 1996; Cooke, 1999). Knowing the true values that generated the datasets is a major advantage when attempting to ascertain the bias, precision, and other characteristics of an estimator. The main burden this methods imposes is the generation of values that contain all the important

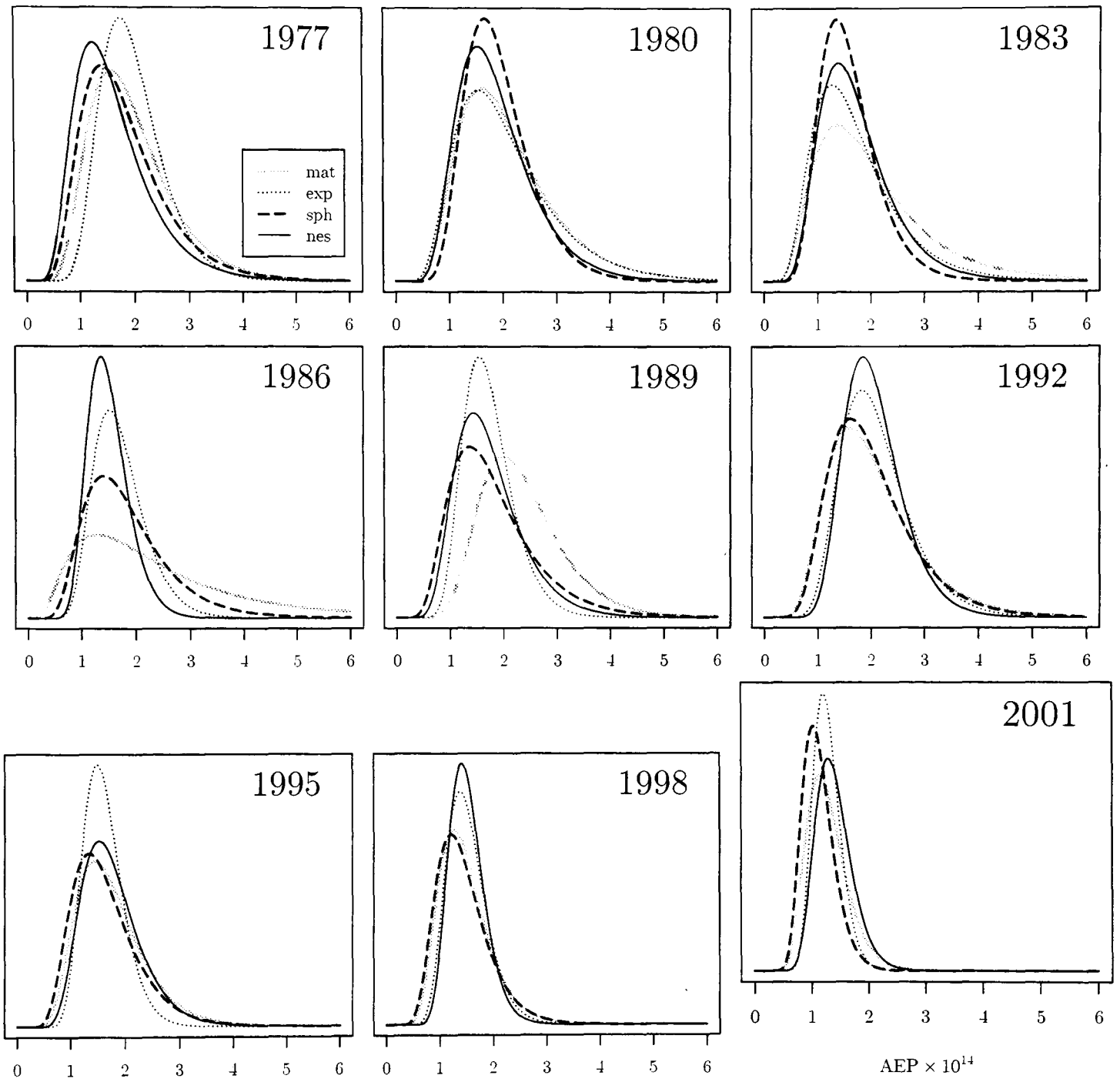


Figure 4.5: Probability distributions of the kriging estimates of annual egg production ($\times 10^{15}$) for every year and model. Plots show parametric smoothed curves of the histograms obtained from 500 iterations.

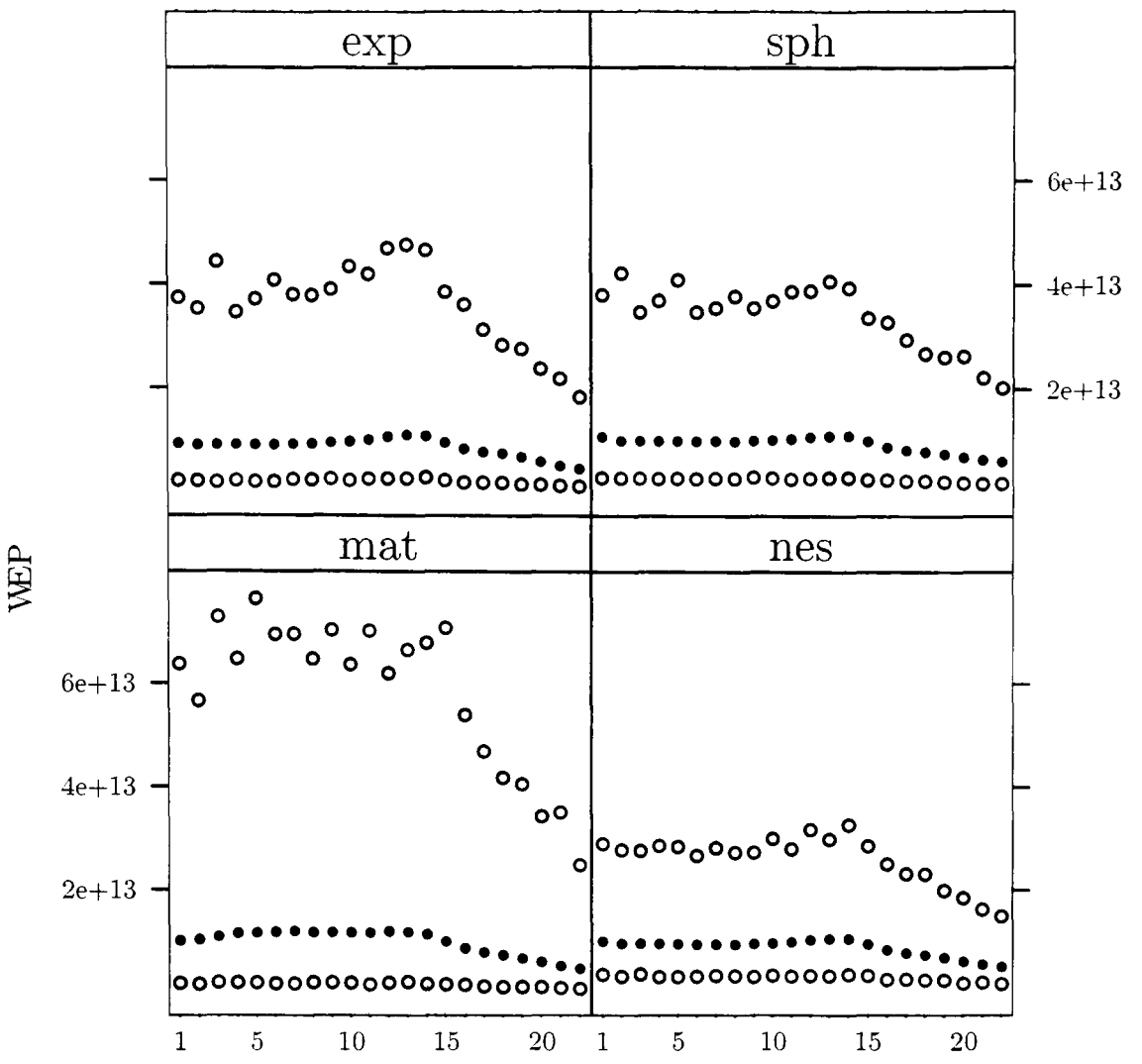


Figure 4.6: Weekly egg production (WEP) estimates for all models and for 1986, shown as median (closed circle) and 95% credibility intervals (open circles).

characteristics of the original data. This can become extremely complex and, depending on the generating mechanism, could unduly influence the results of the whole exercise.

Two simulated datasets were used in this study. The first was generated by Dr D. Beare, at the Marine Laboratory Aberdeen, based on a Generalized Additive Model (GAM) of mackerel egg abundance (Augustin *et al.*, 1998). Weekly surfaces of egg abundance were generated based on the 1998 dataset, with the addition of a negative-binomial random error. The actual 1998 sampling pattern was then applied to the data to generate a possible realisation of that year's mackerel egg survey. The procedure was repeated with subsequent values from the error probability distribution and 1,000 datasets were generated.

A second dataset was generated also based on the 1998 survey. The output of a Gaussian Sequential Simulation (SGS) estimator (Goovaerts, 1997), developed by Dr C. Imrie at Imperial College (Imrie *et al.*, 2003), was used here. The SGS paradigm is based on the generation of several joint realisations of a number, N, of random variables at locations in the study area. These realisations are conditional on the data set under scrutiny. Simulated values at data locations are drawn from a Gaussian probability density function centred on the datum value and with variance equal to the sampling variance (Goovaerts, 1997; Deutsch & Journel, 1998). The algorithm proceeds sequentially by drawing a first

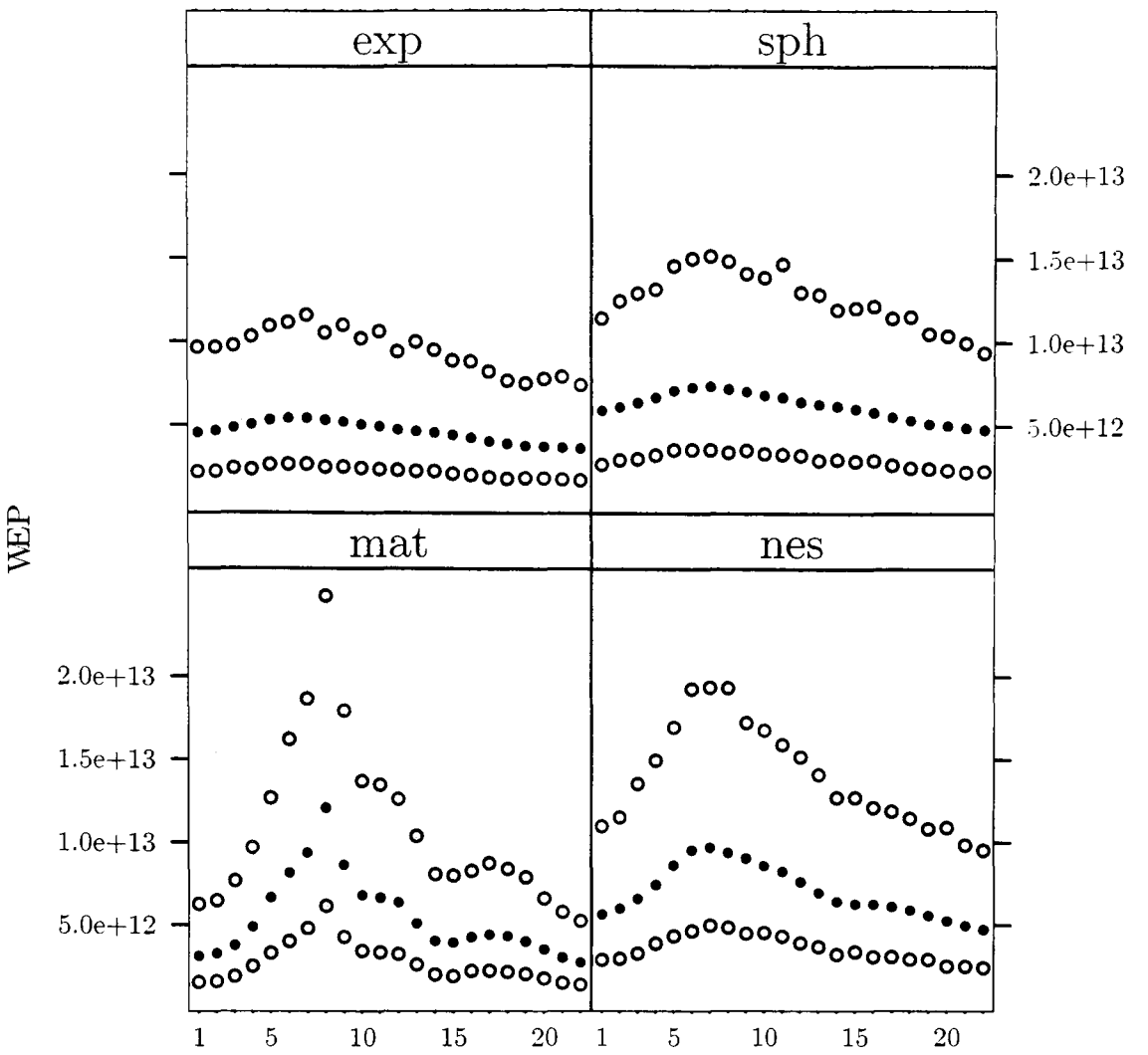


Figure 4.7: Weekly egg production (WEP) estimates for all models and for 1998, shown as median (closed circle) and 95% credibility intervals (open circles).

simulated value conditioned on the N data points, which then becomes a conditioning value for subsequent draws. All nodes in the grid are visited until a complete realisation of the random function is generated. Any number of realisations can then be obtained by repeating the sequential process, usually following different paths to visit the nodes. When using SGS to generate datasets, the variance of the error pdf is increased so as to introduce extra noise in the simulated dataset.

Both datasets were then sampled according to the 1998 survey. No attempt was made at mimicking the expected behaviour of the survey vessels when confronted with the simulated egg abundances. In this behaviour sampling is expected to continue in the western limit of spawning until a sample of zero eggs has occurred. The extension of sampling in this way usually has an impact on what and how many stations are sampled further down the trip, as days at sea are always limited. Replicating this adaptive mechanism in the sampling design when applied to the simulated datasets would have been an extremely complex exercise.

The Bayesian geostatistical estimators were applied to 250 realisations of both simulated datasets. Estimates of total AEP and associated coefficient of variation where compared with those obtained from the traditional (ICES, 1999a) and GAM (Augustin

et al., 1998) estimators, and with the true values used to generate the datasets. No attempt was made to analyse in greater detail the reasons for the differences found among estimators on these datasets, due to constraints in time and availability of the GAM code.

The obvious limitation to the application of simulated datasets is how well the data patterns encountered in nature can be replicated. The models used would always introduce a number of simplifications and assumptions which might make the dataset unacceptably different from reality. Inferences thus made would then have limited applicability in assessing the real behaviour of an estimator. Further problems arise when one of the estimators under scrutiny is itself used to generate the data. This is the case with the GAM model and, to a lesser extent, with SGS.

4.4.2 Model comparison using simulated datasets

Variogram modelling for the two simulated datasets was carried out using the same methods presented in Chapter 3 for the ICES survey data. Models in this case were non-hierarchical, as only one survey was being modelled in each case. Prior probability distributions were defined based on the posterior pdf of the mean hyperparameters from the analysis of ICES survey data. The variograms obtained for both simulated datasets clearly reflect their different originating models (Figure 4.8). Sample variograms for the GAM-generated survey are poorly defined, and consequently the models are highly uncertain. This could be due to the effect the smoothing splines are having on the covariance structure of the data. Better results were obtained when using the SGS datasets, although in contrast with the *real* 1998 variogram (Figure 3.3), uncertainty appears to be higher, probably due to the lack of a hierarchical setting. Bias in the AEP estimates for the four

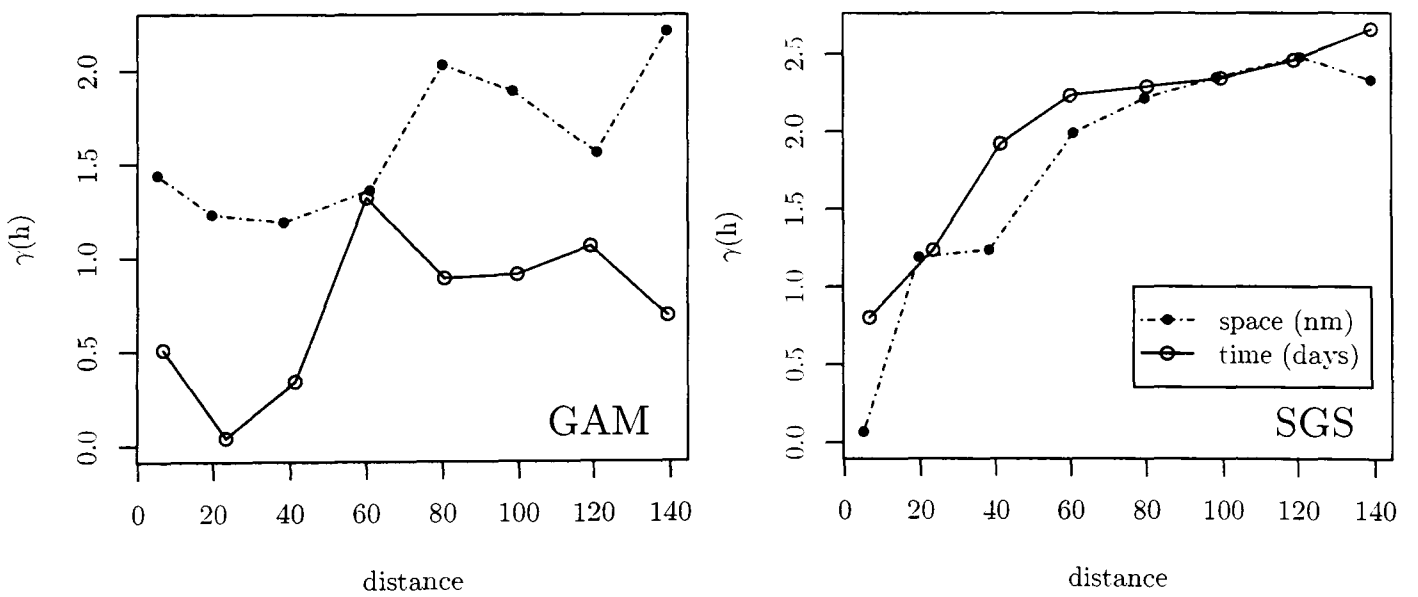


Figure 4.8: Sample variograms (mean of 250 simulations) obtained from the simulated datasets generated from a Generalized Additive Model (GAM) and Sequential Gaussian Simulation (SGS) for 1998.

models, in terms of differences against the "true" simulated values, range from less than

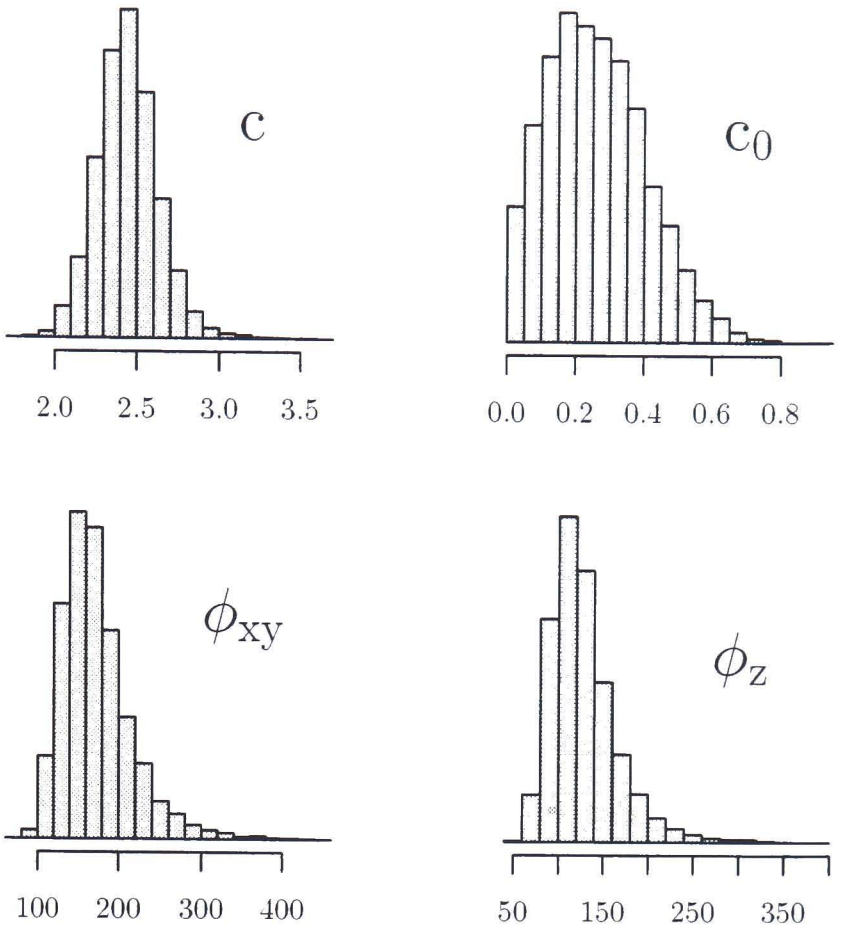


Figure 4.9: Posterior probability distributions of the exponential variogram parameters (sill, nugget and spatial and temporal ranges) obtained using the SGS-generated dataset.

1% up to 8% (Table 4.4) when using the SGS dataset. In contrast, the geostatistical models fare poorly when applied to the GAM generated datasets. Egg production is greatly overestimated by the four models. These results must be considered with caution, as the GAM dataset appears to have lost most of the spatio-temporal structure that is essential for kriging to operate, as can be seen by comparing the sample variogram obtained (Figure 4.8) with those for the original 1998 dataset (Figure 3.3). Despite the limitations

Table 4.4: Median AEP estimates from kriging with the four variogram models using the SGS and GAM-generated simulated datasets

Model		mat	exp	sph	nes	real
SGS	median	1.35	1.28	1.36	1.43	1.42
	difference	-5%	-9.8%	-4%	+1%	
GAM	median	2.12	1.94	2.21	1.87	1.48
	difference	+43%	+31%	+49%	+26%	

of the simulated data outlined above, again the nested (exp+sph) model appears to fare better and provide less biased estimates than the other three.

4.5 Inferences on spawning dynamics from kriging

In addition to its use for estimation of egg production, fine scale spatio-temporal models such as the ones presented here provide interesting opportunities for visualisation and understanding of patterns in spawning at finer scales than those obtained from inspection of the survey data alone (Beare & Reid, 2002). Two possible analyses are presented here. One attempts to reveal the change in location of the centre of activity of spawning along the season. This could be of use during the design phase of future ichthyoplankton surveys. The second relates to the precise timing of spawning, as inferred from detailed maps of weekly egg abundance.

4.5.1 Centre of gravity of spawning

A simple index of changes in the temporal distribution of spawning along the season can be obtained by estimating the centroid, or centre of gravity, of the egg abundance matrices. The centroid of a discrete N-by-M two-dimensional matrix, $A = (a_{ij})$, is calculated as

$$\hat{x} = \sum_{i=1}^{N} \sum_{j=1}^{M} m_{ij} a_{ij} \tag{4.3}$$

where m_{ij} are the normalised matrix values $m_{ij} = a_{ij}/\mathcal{N}$, with $\mathcal{N}$ equal to

$$\mathcal{N} = \sum_{i=1}^{N} \sum_{j=1}^{M} a_{ij} \tag{4.4}$$

x_{ij} is the location vector of each matrix cell (for example $x_{1,2} = [1, 2]$), and a_{ij} are the matrix values at each location. This calculation can be easily extended to a third dimension to estimate the centroid of the whole spatio-temporal array. This metric averages over the whole area or volume, and does not fully reflect changes in spawning distribution. It would, for example, not detect a change such as that created by a situation where spawning activity is split from a central area to two equally distant locations.

An example of changes in the spatial intensity of spawning can be observed in Figure 4.10, which shows the position in latitude and longitude of the spawning centroid by week for 1998. The values were estimated by block kriging based on the nested model and for 100 randomly selected simulations. The different colours show replicates for each week, with the black-closed circles marking the mean value for the x and y position of the centroid by week.

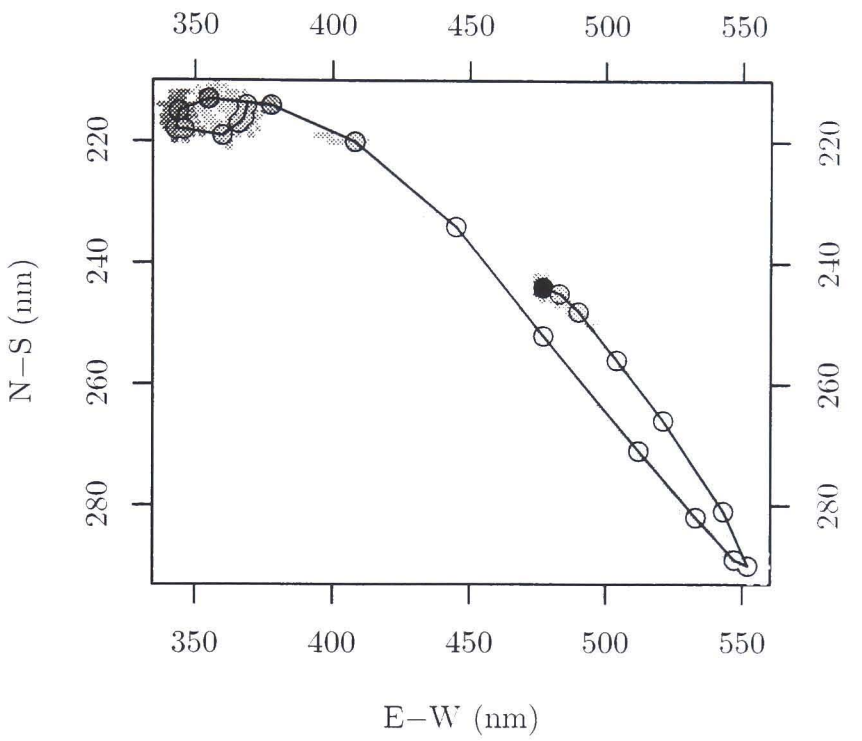


Figure 4.10: Position by week of the centroid of the egg abundance matrix, for 1998 and using the nested model, along the North-South and East-West dimensions and in distances from the assumed (0,0) point at 60° N - 20° W. The black dot indicates week 7, the start of the spawning season. Replicates for each week are denoted by different colours.

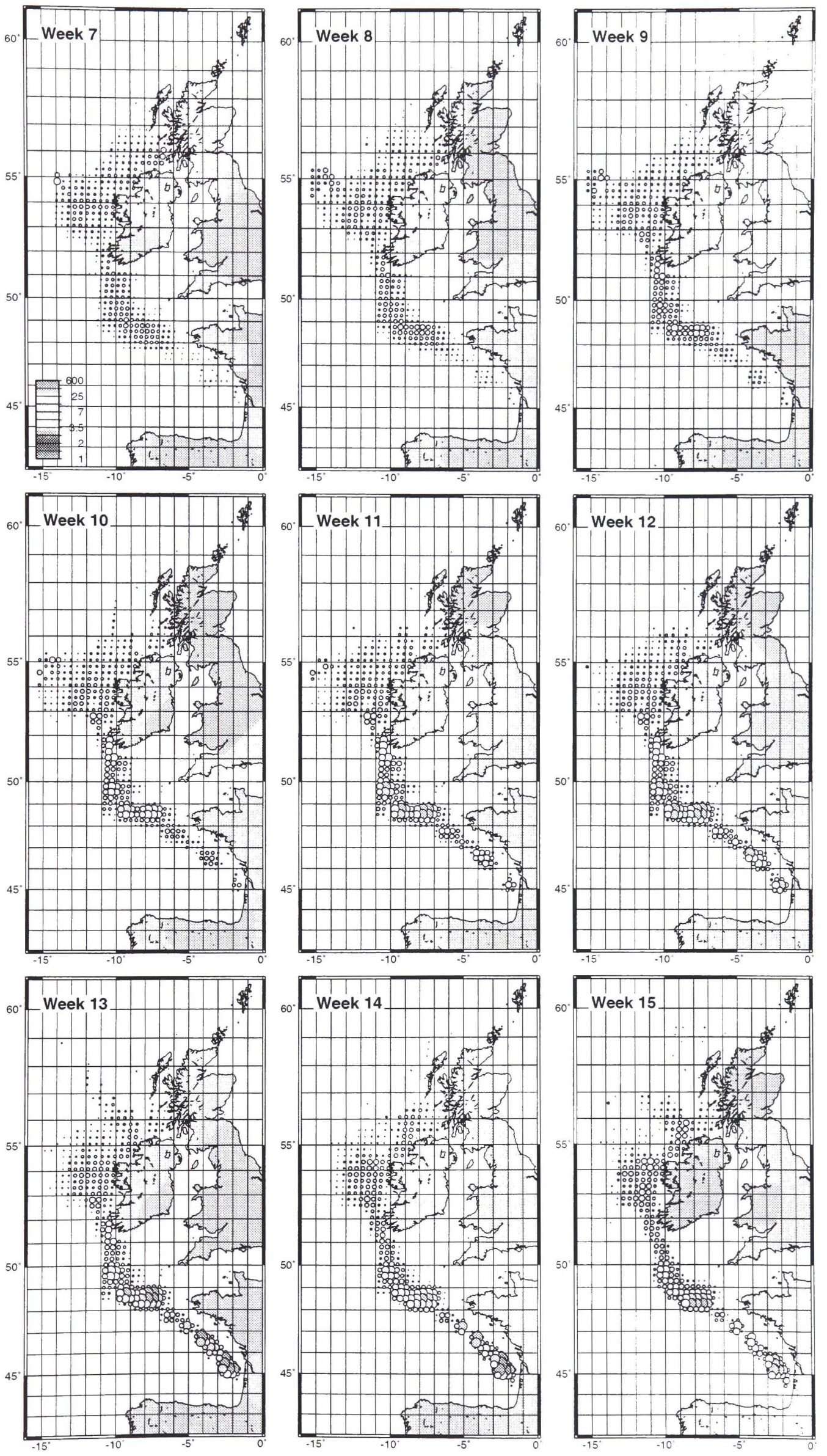


Figure 4.11: Weekly mackerel egg abundances (egg/m^3) in 1998 as estimated by block kriging using the Matérn model. Sizes and colours in log scale.

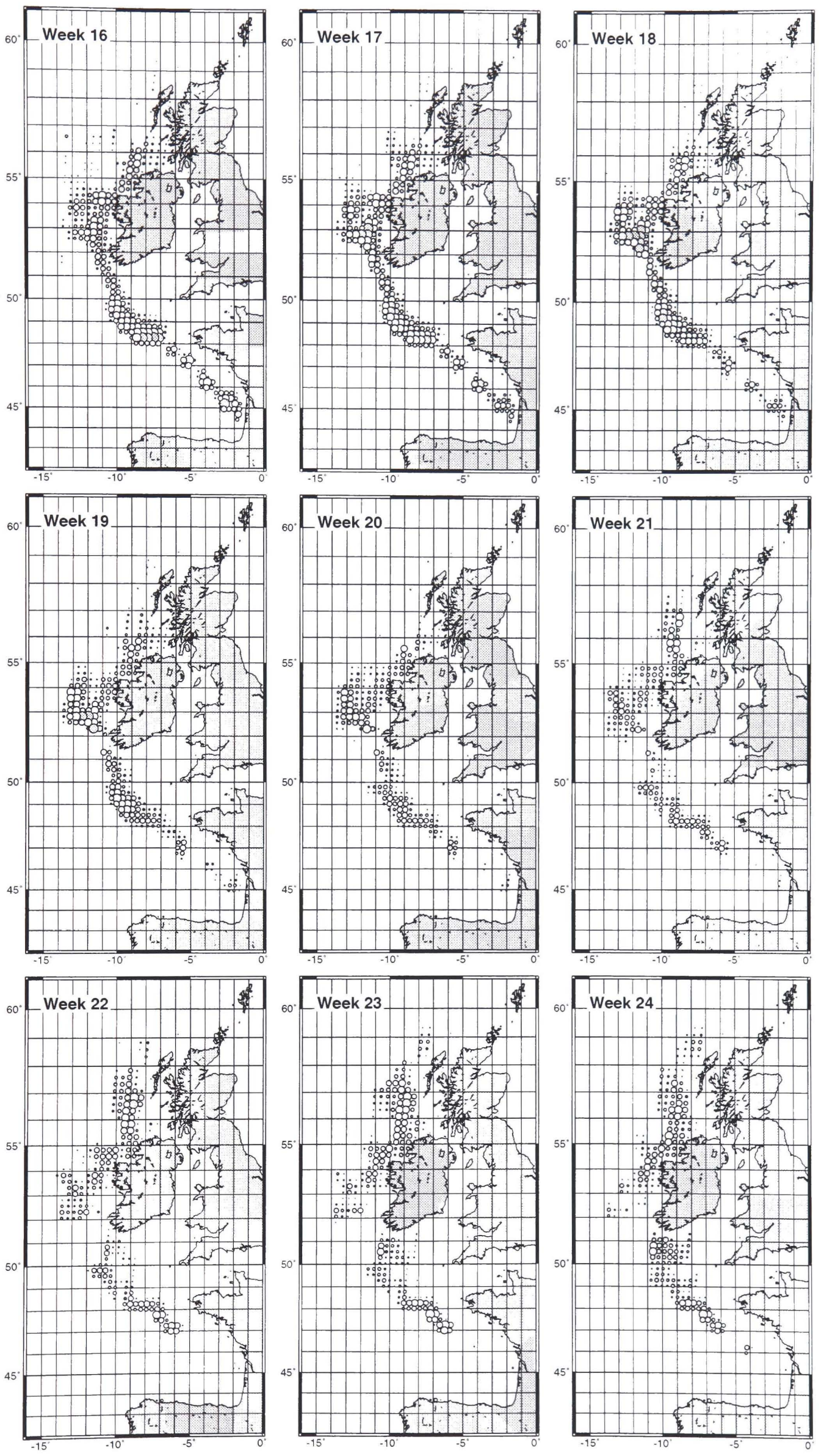


Figure 4.11: Continued

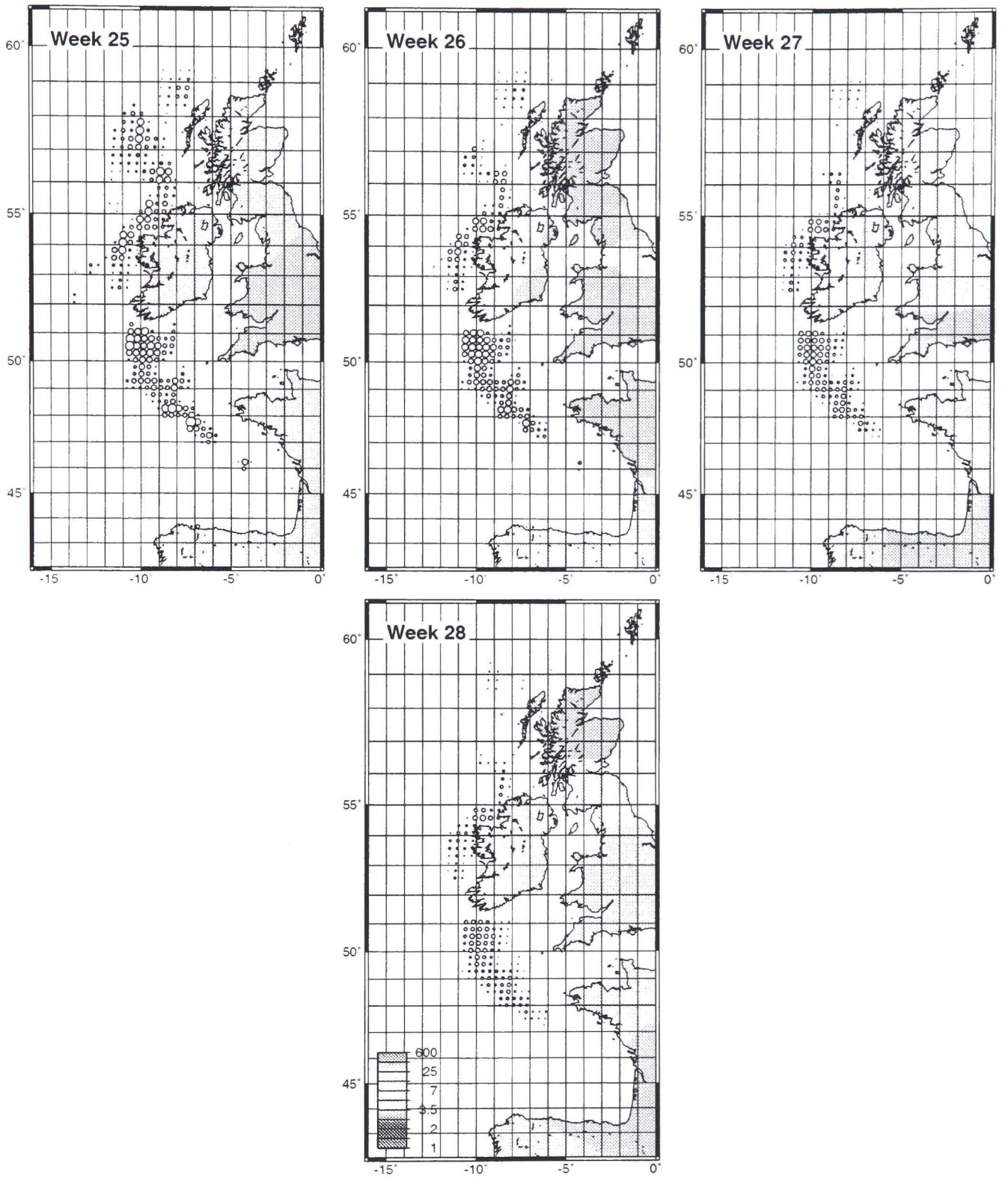


Figure 4.11: Continued

4.5.2 Timing of spawning

Kriging over fine temporal scales provides a valuable tool for analysis of changes in the time pattern of spawning. Detailed investigation of environmental effects on spawning dynamics, for example, can greatly benefit from the weekly estimates of egg abundance provided by the kriging estimator. Detailed maps of mackerel egg spawning by week can also be generated from the kriging results. These also help in visually assessing the results obtained from the kriging estimator.

The maps of estimated weekly egg production for 1998 (Figure 4.11) show a general pattern of spawning distribution by week that is in accordance with previous studies, but also reveal new features that might need further explanation. Egg production in the north of the spawning area appears much earlier than expected (Weeks 7 to 14). This could be due to an excessive generosity of the kriging procedure when estimating across time. It is also interesting to note the separation of spawning in two centres after week 19, one in the Bay of Biscay and one in the west and north of Ireland. This might indicate that spawning activity in the Bay of Biscay remains relatively high even late in the season, when it was assumed to have been displaced further north. The increase in the mean summer sea surface temperature in this area detected over the last few years (I. Sancristobal, AZTI Fundazioa, *pers. com.*) could have an important impact in the temporal and spatial distribution, and even on the total spawn for NEA mackerel, as spawning could potentially extend over longer periods and areas.

4.6 Discussion

Estimates of annual egg production have been produced based on block kriging over hierarchical Bayesian variogram models. Kriging has been carried out in a Monte Carlo procedure in order to introduce the uncertainty in spatio-temporal covariance modelling into the final estimate. The procedure has been applied to ICES datasets of NEA mackerel egg surveys, and the results from this estimator have been compared with those obtained by other methods. The multiple uses of a geostatistical estimator of this kind have been presented, including total egg production estimation and analysis of spatio-temporal dynamics in spawning.

The combination of Bayesian variography and Monte Carlo kriging ensure that all identified sources of uncertainty in the spatio-temporal modelling of pelagic egg production are incorporated in the final variance estimate. In contrast with other geostatistical methods previously used on fisheries data (Sullivan, 1991; Bez & Rivoirard, 2001), estimates presented here do not seem to suffer from over estimation of precision, as models are more parsimonious. Underestimating the variance of an SSB estimate can lead to over-confident risk assessments and over-optimistic management.

In general the results obtained for NEA mackerel mean AEP are not in contradiction with those published elsewhere (Augustin *et al.*, 1998; Imrie *et al.*, 2003; ICES, 2004b), although substantial differences are present. The variability in the estimates appears also to

be within the assumptions or estimates found in those studies. Variability was surprisingly low at the beginning of the series, probably due to the reduced area and time covered. Sampling was confined to a more homogeneous section of the spawning space-time volume, centred around peak spawning areas and periods. Variability increases dramatically in the 1980s, when sampling covers a larger percentage of the spawning volume. Sampling effort increases too, but not as much as coverage does, and data become more sparse in time and space. With the increase in ship time observed in the 1990s, variability appears to decrease, and the last three years of egg surveys analysed (1995, 1998 and 2001) show marked improvements in the precision of the final AEP estimate.

The variance estimates presented in Table 4.2 represent the variability in the mean estimates of egg production, obtained through the Monte Carlo block kriging procedure. Kriging variance is not considered to be accurately estimated when lognormal kriging is applied to transformed data (Goovaerts, 1997). Furthermore, problems were found in the calculation of the kriging variance for some models and in a number of years. The kriging procedure was sometimes unable to provide an estimate of variance probably due to the sparseness of the data. Nevertheless, the key sources of variability can be considered to be incorporated in the analysis. Sampling variability has been introduced in the calculation of the constant of proportionality, q, while estimation variability is accounted for by the combined Bayesian variograms and Monte Carlo kriging procedure.

As explained in previous sections, a number of parameters were not estimated nor was their uncertainty incorporated in the MC procedure. The impact on the obtained estimates of the chosen values for the parameters controlling the kriging algorithm in particular, could have been incorporated explicitly. The associated cost in computing time, however, precluded its proper inclusion. The development of a complete likelihood function for three-dimensional block kriging, encompassing both variogram modelling and kriging estimation, could maybe provide an opportunity for a fuller evaluation of the uncertainty and bias generated by various steps of the kriging procedure. No implementation of such approach has been developed so far, although the directions of some recent developments in spatial statistics (for example Diggle *et al.*, 1998), make the prospect of a coherent Bayesian geostatistical framework appear closer.

An assessment of estimator quality has been limited by the inability of various methods to generate accurate replicates of the original NEA mackerel datasets. On the one hand, the GAM-generated simulated data did not compare in covariance structure to the original, as shown by the sample variograms obtained. This is likely due to the effect of the smoothing splines, controlled in great part by the chosen degrees of freedom allowed to them. It might be the case that increasing the degrees of freedom of the GAM splines, although having a negative effect on the fitting procedure and on its use for AEP estimation, could be beneficial when the GAM is used to generate simulated data. The SGS-generated data were again a by-product of an estimation procedure, and were not designed with data simulation in mind. The similarity of the procedures used to generate

and later analyse the data, both based on variogram modelling, also limit the usefulness of the results obtained. It is expected that a geostatistical procedure should be able to estimate reasonably well the spatio-temporal structure generated based on similar premises. Testing a statistical estimator on simulated datasets has sometimes limited value in terms of assessing its static quality (Restrepo *et al.*, 1990). More useful results could be obtained about the changing behaviour of the estimator under changing conditions in the structure, quality or quantity of the data. The focus would be then shifted from assessing whether the estimator performs well in a supposedly accurate representation of reality, to understanding how well the estimator would perform if conditions were to change in certain directions. In the case of fisheries assessment, an important issue to consider is how well would the estimator detect decreases in spawning stock biomass beyond safe limits or quantify the risks of collapse or damage to the stock.

It is nevertheless possible to assert that, of the four variogram models tested on the NEA mackerel egg surveys dataset, the nested model, combining an exponential and a spherical component, did show the best performance and flexibility. The range of variograms observed, among years and for both dimensions, could be well accomodated with this model. The increased number of parameters did not have a negative impact in the McMC procedure. Measures of model quality presented in the previous chapter also appeared to select this model over the other three. The results presented in this chapter, especially when analysed at a fine temporal level, further corroborated this finding. Although this is obviously of importance for the NEA mackerel dataset only, it also shows how difficult the balance between model complexity, fit to data, and usability for prediction can be in many cases. Furthermore when applying hierarchical models that attempt to uncover the common patterns over a number of surveys.

The estimates of annual egg production now obtained from Bayesian hierarchical variogram models and Monte Carlo three-dimensional block kriging, will generate, when combined with estimates of fecundity, final estimates of spawning stock biomass. Prior to that, the AEP estimates will be corrected by the application of the bias and error model presented in Chapter 2. Combining in this way all identified sources of variability will allow the exploration of a range of scenarios on the present state of the stock. This will be presented over the next chapter based on the estimates obtained from the best model.

Chapter 5

Comparison of estimators and implications for assessment and management

The previous chapters have first presented a detailed analyses of the procedure carried out for the estimation of Annual Egg Production. A quantitative estimate of bias and error, in the form of a probabilistic constant of proportionality (q), was then developed based on the knowledge gathered. Then, alternative estimators of egg production based on kriging over Bayesian hierarchical models of spatio-temporal covariance were built, tested and compared to previous estimates of AEP from the NEA mackerel egg survey data.

Estimates of annual egg production do not provide per se with sufficient information for assessment and management purposes. They need to be coupled with estimates of realised fecundity, linking in this way egg production to the spawning stock biomass that generated them. The uncertainty and possible bias in the procedure used to estimate fecundity in determinate spawners such as mackerel, was reviewed in Chapter 2, and the results incorporated to the estimated constant q.

Kriging around the estimated variogram models was carried out for the four variogram hierarchical models. The most flexible model, combining a short-range exponential and a long-range spherical model, appeared to outperform the other two and has been thus chosen for the analyses presented in this last chapter.

Both sections of this work will now be brought together, in order to generate probabilistic estimates of SSB. These will then be fed to the assessment procedure presently in use for NEA mackerel to provide an indication of how uncertain the estimates of SSB and the management reference points might be when considering all the error and bias present in the AEP estimate, and how changes in the methodology for estimating AEP affect our perception of the status of the stock. The implications for management of the NEA mackerel stock of the findings of this study will be further elaborated. The newly generated

estimates of SSB will be introduced into the currently applied assessment method in order to highlight the different outlook and status of the stock obtained with them. The question of whether SSB estimates obtained from egg surveys should be considered as relative or absolute indices of stock size is discussed with reference to the impact of that choice on NEA mackerel assessment.

Finally, a discussion of the results and conclusions that can be extracted from the present study is undertaken. Consideration is given to the possible shortcomings of the approach developed here, future lines of work that could correct those, and what other areas of research might be of interest in this area.

5.1 Bias-corrected estimates of annual egg production

The constant of proportionality, q, generated in Chapter 2 from the application of the bias model to every survey year dataset, has now to be combined with the AEP estimates obtained through kriging in Chapter 4. We would arrive in this way at a bias-corrected AEP estimate that also incorporates the error introduced along the sampling procedure (Table 5.1).

Table 5.1: Median, standard deviation, coefficient of variation, and ratio over the standard estimates, of the bias-corrected estimates of annual egg production for each year of the nested (exp+sph) model, in eggs $\times 10^{15}$, and mean estaimte of the Traditional AEP estimator.

	Nes				Trad
	median	sd	cv	ratio	mean
1977	1.53	0.74	0.48	1.64	1.98
1980	1.75	1.20	0.67	2.26	1.84
1983	1.53	0.72	0.44	1.74	1.53
1986	1.49	1.09	0.68	1.78	1.24
1989	1.63	1.27	0.66	1.83	1.32
1992	2.11	1.72	0.75	2.33	1.94
1995	1.57	0.48	0.33	1.81	1.49
1998	1.50	0.46	0.29	1.52	1.37
2001	1.40	0.53	0.36	1.41	1.21

The bias model that is now applied to the kriging estimates of AEP takes into account the identified sources of bias and uncertainty in the egg sampling procedure. Bias is introduced first by the decision to limit sampling to a depth of 200 m, and secondly by the assumption that temperature at 20 m represents accurately the temperature experienced by eggs developing along the water column. The bias model was applied to each year of data, taking into account the number of samples taken in waters deeper than 200 m and before or after 120 days from the 1st of January, where the effects of temperature stratification roughly start affecting the distribution of eggs in the water column (see

diagram in page 60). Although in absolute terms both sources of bias appear to be very small, and tend to cancel each other, they were still introduced in the generation of yearly values for q, the constant of proportionality. This constant also included the uncertainty in raw egg counts introduced by errors in egg counting and staging, and in the estimation through a flowmeter of the volume of water filtered by the egg samplers.

Once the bias-model has been applied to all years of survey data, and the quantified sampling uncertainty incorporated, the bias-corrected estimates do not differ greatly from those presented in the previous chapter (Table 4.2, page 107), as can be seen from Table 5.1. The main reason, as discussed when the bias model was laid out, is that samples in the time and area window where most bias might be introduced are few and in general show low values of egg abundance. This is even more marked the further back we go in the survey time series. Also, the tendency of both sources of bias to cancel each other out means bias-corrected estimates are not too different from the original ones in their central tendencies.

A graphical representation of the AEP estimates and the corresponding 95% credibility intervals shows how variable the estimation of egg production in some years are. Two of the surveys which showed remarkably flat variograms, 1986 and 1989, present wider credibility intervals. This probably shows the effect of a not completely appropiate variogram model, when applied to those years. Although the median estimates are within the values that could be expected, the final estimated variability appears to be relatively high. The increase in sampling effort is reflected in a marked decrease of the estimate variance for the last three years of the survey, 1995-2001. Further comparison of the final results among the four models tested can be obtained by inspection of the bias-corrected annual egg production probability distributions (Figure 5.2).

5.2 Estimates of spawning stock biomass from kriging-derived annual egg production

Estimates of SSB are presented here for all years. Fecundity estimates from 1977 to 1989 were not available in the same detailed form (tables of potential fecundity, atresia and weight per female sampled) as for the 1992-2001 period. Mean realised fecundity for those years was reconstructed from ICES Working Group reports (ICES, 1979, 1981, 1984, 1987, 1990). Estimates provided as a non-linear relationship between length and fecundity were transformed to weight-fecundity by using available growth curves (ICES, 2004b) and estimates of mean weight in the catch. The mean-variance relationship estimated for the 1992-2001 fecundity data was also applied to the previous years, despite indications that variance was likely to be higher at the start of the survey series. Fecundity estimation in the earlier years of the survey did not take into account the influence of atresia (ICES, 1981), which only started to be estimated for the 1989 survey (ICES, 1990). No information was available that could indicate how variable fecundity estimates from 1977-1989 might be.

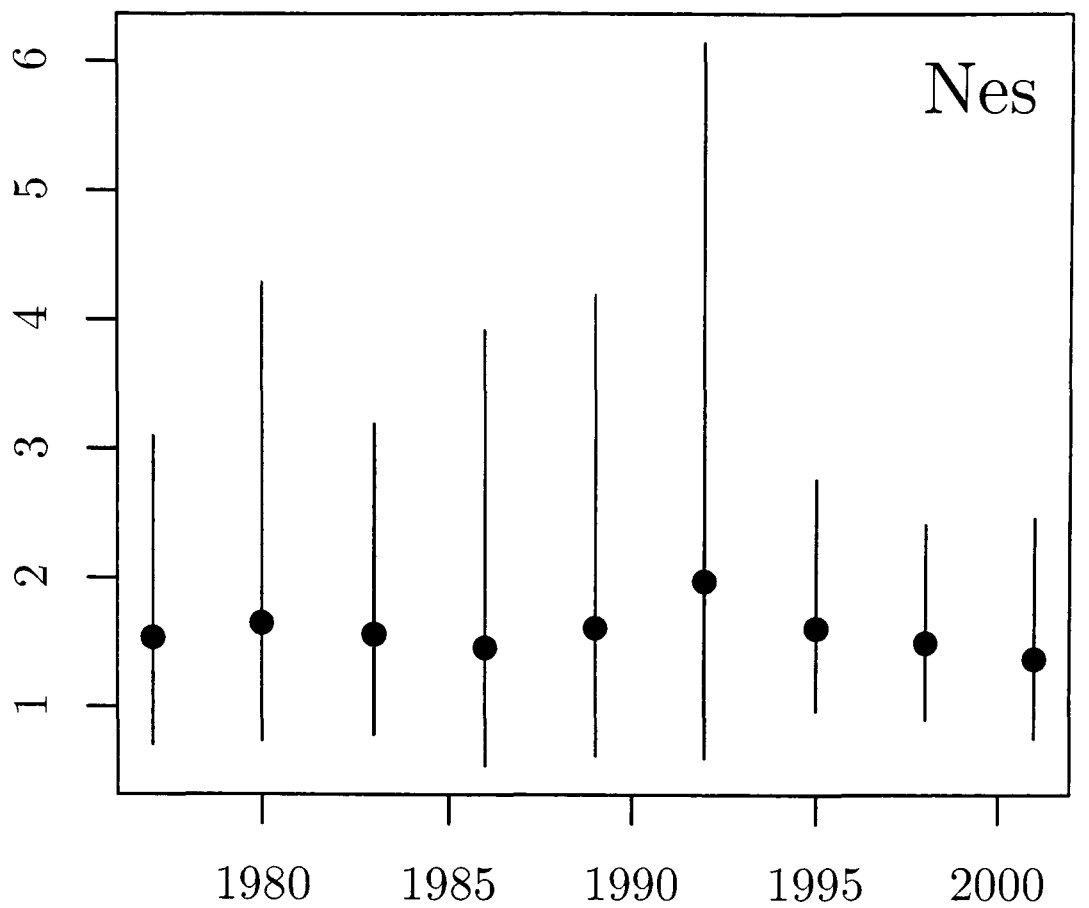


Figure 5.1: Median estimates of bias-corrected annual egg production and 95% credibility intervals for the nested model.

The probability distributions of spawning stock biomass for all surveys are presented in Figure 5.3. These were obtained by combining the 500 estimates obtained by kriging with 500 random samples from the probability distributions of realised fecundity (see page 46). Except for 1998, results provide a fairly similar picture to that observed in the AEP values. Variability, represented here by the 95% credibility intervals, appears to be lower at the start of the series, increasing in the 1980s as sampling was extended, only to decrease as sampling became more intense and spawning limits were better sampled with the adoption of adaptive sampling (Figure 5.4).

The amount is variability uncovered in these results appears to be much larger that the amount of variation in the median estimates of SSB along the time series. The noise to signal ratio, therefore, could prove to be a difficulty when attempting to discern a trend in stock biomass if the index variability were to be included in the stock assessment procedure.

5.3 Factors not included in the analyses

Despite the apparent simplicity of most egg production methods, involving only egg production and fecundity, arriving at an estimate of SSB from a number of egg counts is in reality a complex procedure. As explained in Chapter 2, the many steps and procedures involved can all have an important influence in the final estimate. Despite the detailed analyses of the AEP methods attempted here, a number of factors could not be included, and their relative influence in the uncertainty of the estimation procedure could not be

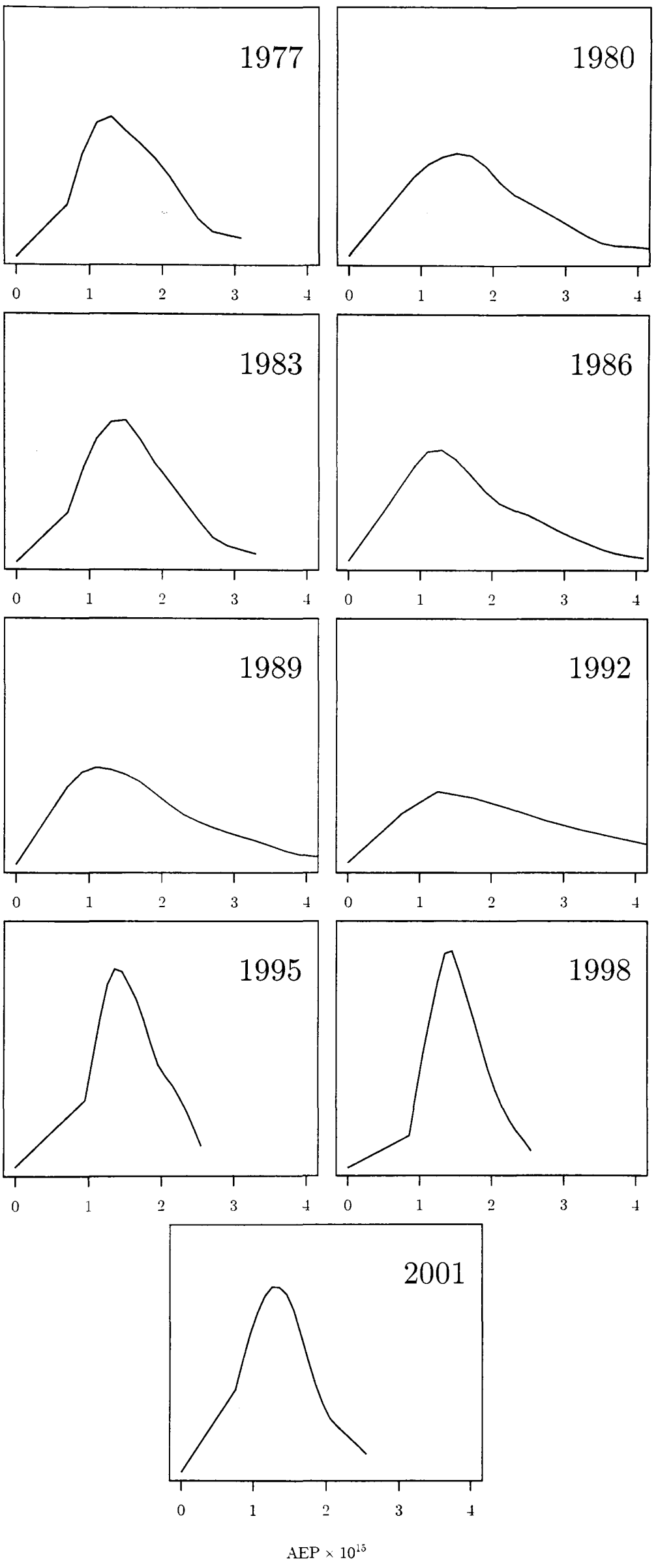


Figure 5.2: Probability distributions of the bias-corrected annual egg production for the nested variogram model.

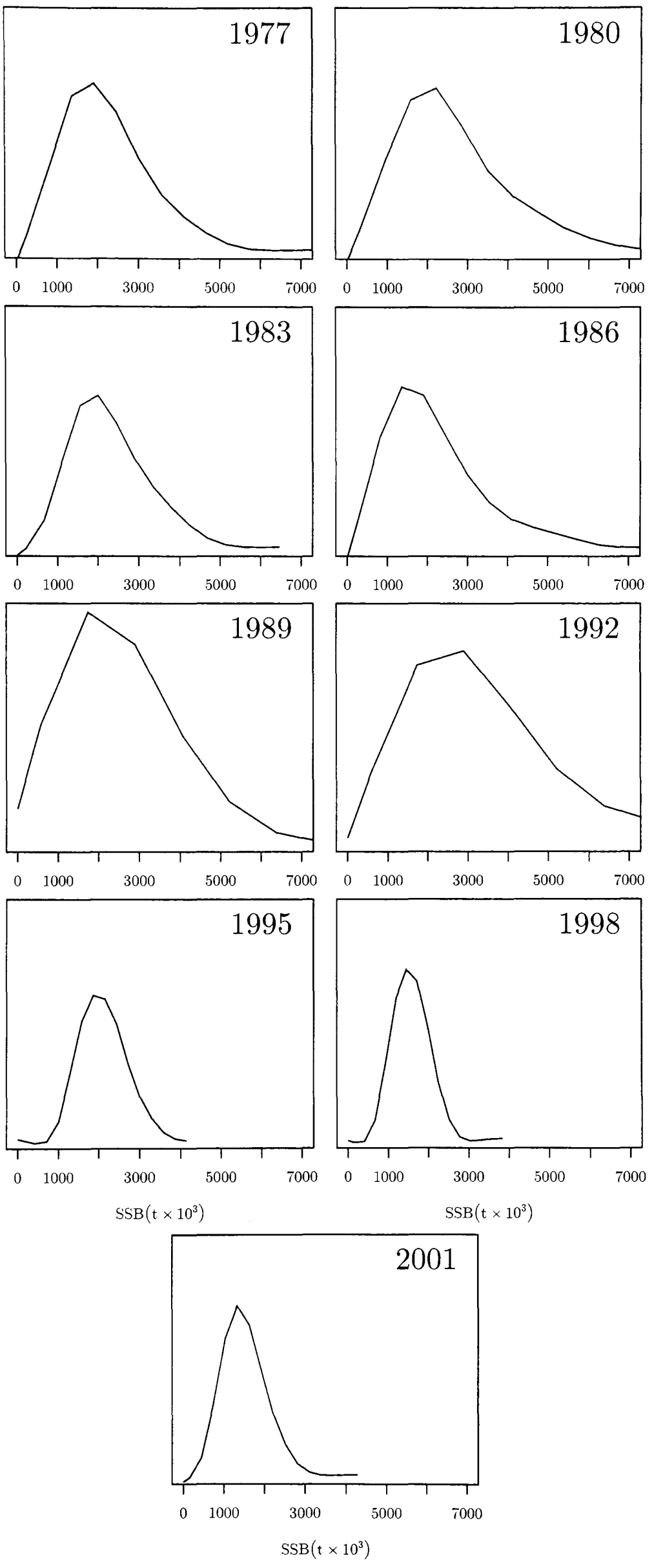


Figure 5.3: Probability distributions of spawning stock biomass obtained from the nested model

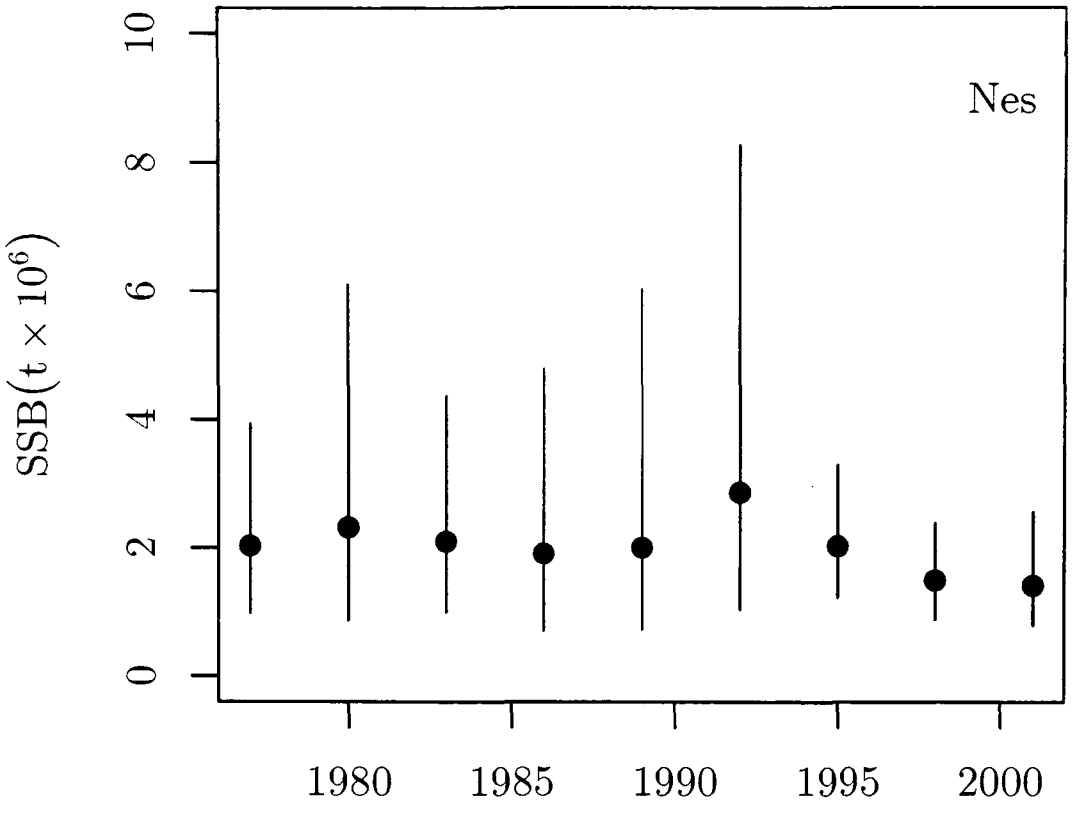


Figure 5.4: Median spawning stock biomass estimate and 95% credibility intervals from kriging.

quantified. The most important factors not included in this analysis are now briefly reviewed, and their possible influence, in qualitative terms, is discussed.

Egg mortality

A number of the applications of the AEP method presented in Table 1.1, make use of the abundance estimates obtained for eggs at all development stages. In contrast, the ICES NEA mackerel survey considers the abundance of the younger, stage I, eggs only (ICES, 1999a). This is due to the known difficulty in correctly identifying the development stage of older eggs of mackerel (ICES, 2001). The first obvious impact of this decision is the loss of valuable information, as a considerable section of the population under study, that of pelagic eggs older than stage I, is not incorporated in the analysis. Sampling occurring in an area where an important spawning event has taken place, but with enough delay to allow eggs to proceed to stage II or further, for example, could still be detected from counts of older eggs.

As mentioned in the presentation of the AEP method (section 2.3, p. 41), use of data on all egg stages depends on an estimate of egg mortality. This can be generated from the egg abundances themselves, or be available from other studies. Ideally this should include how mortality is likely to vary with date, area and oceanographic conditions, such as wind dispersion, temperature and mixing. At present there are only some indications of the probable mortality rate of NEA mackerel eggs, tentatively considered from some research to lie at around 8% per stage (S. Coombs, *pers. com*). Recent experiments on repeated sampling of a single spawning stop for the whole of the duration of stage I, has estimated a much larger value for stage I mortality rates whih could mean egg abundance

at stage I has been understimated by between 50 and 300% (D. Beare, *pers. com*). The present difficulties in mackerel egg staging preclude the estimation of the mortality rate directly from egg counts of all stages (as shown in equation 2.5), and its use in the AEP estimation procedure. However, advances in egg staging by molecular and image analysis methods could enable the use of the whole egg datasets in the near future.

Even when limiting the estimation procedure to the first stage, the question of natural mortality remains unaccounted for. Eggs at stage I of their development are bound to be subject to natural mortality for the duration of this stage. Pelagic eggs are in the size range that makes them susceptible to predation by a number of marine species, including mackerel itself (Lockwood, 1988). And mortality is likely to depend on the length of the development process, itself related to temperature. Consequently, a rough estimate of mortality rate per development stage does not allow for detailed analysis of the effect of ignoring mortality on the final AEP estimate by year. Nevertheless, estimates based on stage I abundance without incorporating natural mortality are bound to underestimate, to a greater or lesser extent, the real egg production.

Kriging parameters

A number of parameters that control the kriging procedure were selected *a priori* and fixed in the analysis. This goes clearly against the desired attributes of a Bayesian analysis, where uncertainty at all stages should be explicitly incorporated. Realistically, estimation of parameters like search range, maximum and minimum number of points to consider, and blocksize, could not be incorporated in the kriging procedure applied here. An evaluation of their influence would have had a large cost in computational terms. Their impact might perhaps be evaluated in the future on a limited dataset, or for a single variogram model. The question nevertheless remains on the proper procedure to apply if the parameter values chosen in the present analyses were shown to unduly affect the final results.

5.3.1 Survey design and estimator choice

An important factor when analysing fisheries survey data is the survey design that generated it. Survey design is a well established field with important links to ecological methodology (Krebs, 1989), and general statistical analysis. The design of trawling and acoustic surveys has always been under intense scrutiny (Gunderson, 1993). In contrast, ichthyoplankton surveys have generally been based on a limited set of sampling strategies, perfectly suitable for the standard abundance estimation methods, but that might prove limited or inadequate for AEP estimation using geostatistics, GAMs or other complex methods.

The present study has not covered the issue of sampling design or its influence on egg production estimation in great detail. Despite this, a number of general observations can be made with respect to the interaction between the current sampling design in ICES NEA mackerel egg surveys and the geostatistical methods presented here. The present strategy attempts to cover as intensely as possible the spatial dimension during

a number of short periods of time along the spawning season (ICES. 1999b). This is consistent with the differential treatment of the temporal and spatial dimensions by the *traditional* AEP estimator. More sophisticated methods of spatio-temporal interpolation are not dependent on such a sampling design, and could even benefit from sampling being distributed more evenly along the spatial and temporal dimensions. The capacity of GAMs or geostatistics for accurate interpolation across unsampled rectangles could be put to better use if sampling was distributed more homogeneously on the spatial and temporal dimensions. The present intensity of sampling in space is probably excessive for those estimators, with the temporal gaps proving difficult to model.

5.4 Effect of uncertainty on stock assessment

To further evaluate the implications of the biomass estimates obtained from the present analysis, we will replicate the assessment procedure currently in place in ICES Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy (ICES, 2004b). The SSB estimates obtained from egg surveys are introduced in the assessment procedure in a similar way as other indices of abundance, commonly obtained from catch-per-unit-effort, are. The main difference is that fishery-independent indices are often considered to provide absolute estimates of SSB, thus strongly driving the fitting procedure. The method presently applied in ICES, Integrated Catch Analysis (Patterson & Melvin, 1996), is one of a number of methods that allow an integrated analysis of fishery-independent and catch data. Among other choices, ICA allows the user to specify whether the abundance index from egg surveys is to be considered as an absolute or relative measure of SSB. The objective function will attempt to minimise the residuals obtained when fitting the estimated SSB to both the catch and egg survey data.

Table 5.2: Estimates of spawning stock biomass (in $t \times 10^3$) for the last five years obtained from ICA (Patterson & Melvin, 1996) using the kriging-based estimates of SSB, either as absolute or relative indices of abundance, and considering the whole survey series (1977-2004) or the years used by the working group (1992-2004). The egg survey-based estimate of SSB for 2004 was set equal to 2001, as done by the ICES assessment working group (ICES, 2004b).

	Absolute		Relative	
	1977	1992	1977	1992
1999	1975	1975	2214	2134
2000	1754	1754	2025	1933
2001	1649	1649	1992	1875
2002	1190	1190	1556	1431
2003	987	987	1448	1292

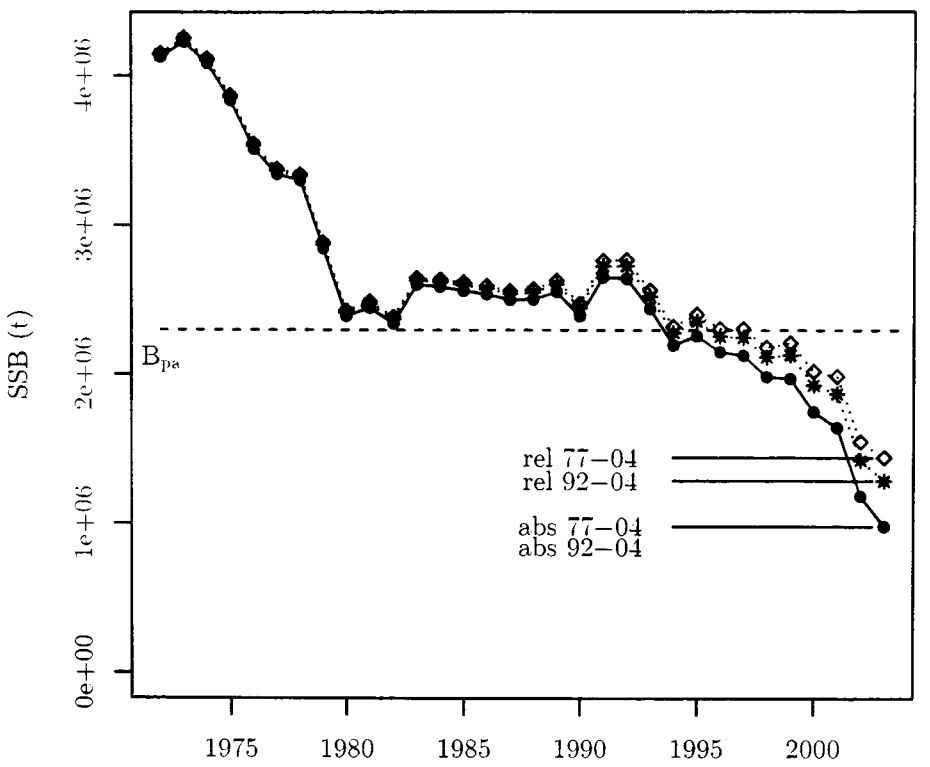


Figure 5.5: Spawning stock biomass estimated by applying ICA (Patterson & Melvin, 1996) using the kriging-based index of abundance either as relative (rel) or absolute (abs), and using the whole survey series (1977-2004) or only the years used by the ICES WGMHSA (1992-2004). Lines for both year options using the index as an absolute estimate show no difference.

5.4.1 Assessments conducted by ICES Working Group

The ICES working group in charge of mackerel assessment (WGMHSA) has been making use of the information on the size of the stock generated by the egg surveys since the start of the series (ICES, 1979). In the latest assessment (ICES, 2004b), the egg survey-based index was considered to provide an absolute estimate of SSB, but only the surveys from 1992 onwards, the most reliable ones in the view of the group, were introduced in the assessment procedure. The latest estimates of stock size obtained during the WGMHSA assessment show a stock in apparent stability since the 1980s, with small fluctuations around a mean for the period 1980-2003 of around 2,665,000 t (Figure 5.6). The disagreement between both time series (catch and surveys) is clearly apparent, especially for 1998. A number of reasons could explain this, the most significant being the possible existence of underreporting and/or discarding in large quantities. Another one is, of course, that the egg surveys are producing overestimates of SSB, at least in some years.

This assessment has been later reformulated by considering the egg survey index as providing a relative index of abundance only (ICES, 2004a). The result of this decision has been an important change in the estimates of SSB and especially on the present situation of the stock relative to the established reference point, B_{pa} (Figure 5.7).

5.4.2 Absolute or relative indices of SSB

A common source of disagreement and uncertainty when fishery-independent surveys of resource abundance are incorporated into the assessment procedure is whether those in-

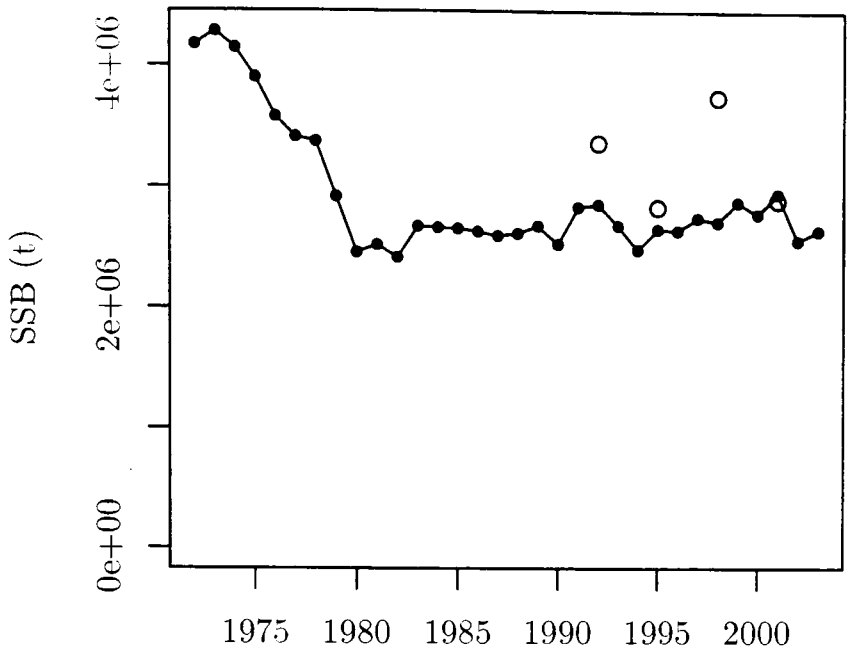


Figure 5.6: Estimates of SSB obtained by ICES Working Group (WGMHSA) employing ICA (Patterson & Melvin, 1996) and considering the egg surveys from 1992 onwards as providing absolute estimates of SSB. The white circles represent the egg survey-derived estimates of SSB, and the line shows the ICA estimate of SSB over time.

dices should be treated as absolute estimates of SSB, or simply as relative abundances indices, similar to those obtained from catch-per-unit-effort of the fishing fleets. The impact on the outlook of the stock can be important if catch data and surveys disagree on the level and dynamics of the stock. For example, the most recent assessment of NEA mackerel undertaken under ICES auspices has provided with two very different impressions on the status of the stock: 3,692,410 t vs. 1,854,720 t SSB for 2003 (ICES, 2004b,a). As surveys of this stock are only carried out every three years – and no other abundance index is available – altering the way the index is considered results in large revisions of the stock size, F and SSB estimates for the whole time series, and consequently catch predictions and Total Allowable Catch (TAC) advice.

The time series of estimated biomass obtained by ICES WGMHSA by considering the AEP method to provide absolute estimates of SSB, can now be compared with that obtained by considering them as relative indices of abundance (Figure 5.7). That was exactly what ICES ACFM did, replicating in every other detail the working group assessment run (ICES, 2004a). Both series diverge markedly from 1995 onwards, and provide very different views on the status and recent history of the stock. The biomass level that complies with precautionary approach has been set by ICES at 2.3 million t (dotted line in Figure 5.7). According to the more pessimistic assessment, this limit has been reached a few times in past, and the stock is now well below B_{pa}.

The way assessments are influenced by the consideration of absolute or relative of fisheries-independent indices depends greatly on how the assessment method utilises this information and combines it with that provided by catch and landings data. In the case of ICA, the objective function tries to minimise the residuals from both the VPA and the

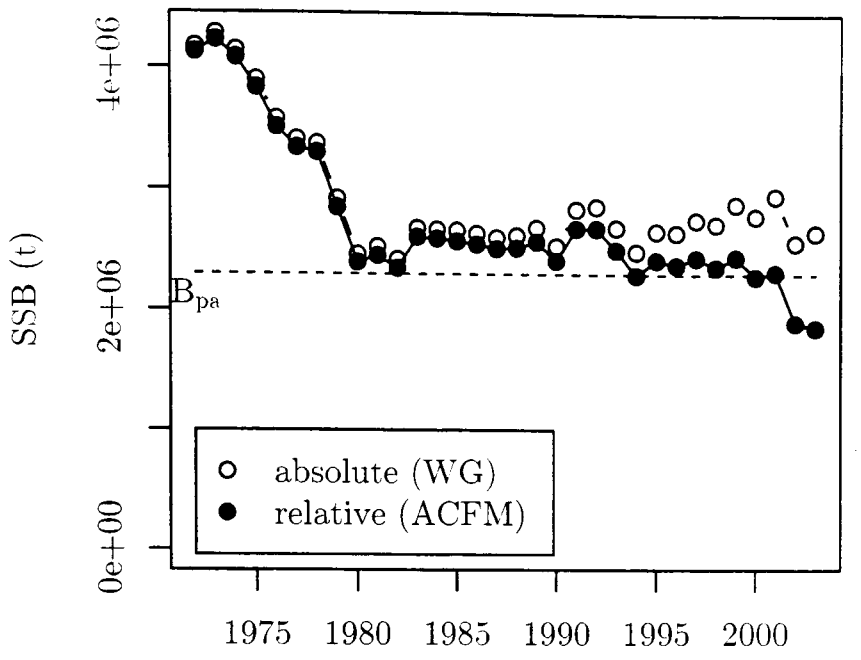


Figure 5.7: SSB trajectories obtained from ICA by considering the egg survey-derived estimate of biomass as relative or absolute. From ICES (2004a).

survey estimates of SSB. The relative importance of each of them in the procedure depends on the yearly weights, which can be assigned or estimated (Patterson & Melvin, 1996). If the later option is chosen, those weights are estimated as the ratio of the variances of both indices, VPA-constructed and survey. This again points to the importance of accurately estimating the variance of the egg survey abundance index.

5.5 Final discussion

A novel procedure for analysing ichthyoplankton survey data and generating probabilistic estimates of spawning stock biomass has been presented here. A review of the survey methodology employed when assessing stocks through the Annual Egg Production Method was carried out. This revision used the historical dataset of egg surveys carried out every three years in NE Atlantic waters by ICES. A bias and sampling error model was constructed that generated a probability distribution for a constant of proportionality, linking estimated and real egg abundance according to the sources of error and bias identified.

An improvement on the spatio-temporal interpolation procedure presently employed in the analysis of NEA mackerel egg survey data was proposed. The method presented is based on geostatistical modelling of the covariance structure in space and time. A Bayesian formulation of spatio-temporal semivariogram models was developed that incorporated the uncertainty in parameter estimation. A number of models were tested to compare the effect of structural uncertainty in variogram modelling. The results of this modelling exercise were applied to the task of estimating total egg abundance from survey data. This was achieved through block kriging over the full spatio-temporal volume. A Monte Carlo procedure was employed to incorporate into this step the uncertainty in variogram modelling quantified before. Finally, estimates of spawning stock biomass were generated from egg production and fecundity to generate probability distributions of SSB.

These were then used in a complete stock assessment and results were compared to those presently accepted for management and advice purposes.

The methodology presented here provides a consistent, if complicated, estimate of variability in annual egg production estimates. It does show considering the variability introduced in the estimate by sampling, through an error model and a constant of proportionality, and by modelling, by hierarchical Bayesian modelling of the covariance structure. The relative impact of model structure was investigated through comparisons of the results obtained by a number of different variogram models. The level of variability uncovered by this analysis is consistent with what can be expected of such a complex sampling exercise, and can be compared in general terms with what other studies have estimated (Simmonds *et al.*, 2003).

5.5.1 Estimator performance

The Bayesian-Monte Carlo geostatistical estimator developed in the previous chapters appears to perform reasonably well on the NEA mackerel egg surveys dataset. Estimates of annual egg production and spawning stock biomass do not seem to be unreasonable when compared to other estimation methods. A complete evaluation of estimator performance has not been possible due to the difficulties in generating simulated datasets that accurately capture the spatio-temporal structure of the original data.

A limitation of the geostatistical methods presented here is the computing power and time necessary. As previously mentioned, the analyses included in this study had to be deployed in a relatively powerful computer cluster and took around 2000 hours of processing time to complete. Application of these methods to future surveys, like the one recently finalised in 2004, will not, however, require special computing facilities. A new run of the hierarchical variogram model can be carried out on a modest computer in less than 10 minutes, and 500 iterations of the kriging procedure for the new data and on a single variogram model can be completed in one hour. In contrast, bootstrapping of the GAM estimator can take up to several days to complete (Beare & Reid, 2002).

Consideration must be taken of the influence of survey design on estimator performance. The present sampling strategy in place for NEA mackerel, shared in general terms with other applications of the AEPM, is based on the separate analysis of temporal and spatial trends. Although no detailed investigation could be conducted on this, spatio-temporal estimators of the type proposed here could be better applied to data collected according to a different survey design, where the spatial and temporal dimensions were sampled more homogeneously. Development of methods for generating appropriate simulated mackerel egg abundance data could open the door to a detailed analysis of alternative sampling design and their possible benefits when combined with spatio-temporal estimators. Limitations in ship time need to be considered, and although a survey with more distributed coverage could theoretically be achieved with the same overall ship time, difficulties could arise with the practical organisation of time at sea. Research vessels can

not be expected to be out at sea for many very short periods, possibly of only a few days, along the spawning season, as this could seriously limit their operational efficiency.

5.5.2 *Management implications*

A fundamental question when assessing fishery-independent indices of abundance is whether they provide absolute or relative estimates of SSB. Considering survey indices in either form generally affects greatly the result of integrated assessments, as has been shown above. Justification for such decisions should ideally be based upon detailed considerations of the statistical qualities of the index.

The results obtained with ICA using the kriging-based series of estimates of SSB differ markedly from those reported by both the ICES WGMHSA and ACFM. Stock biomass might have been below the only biomass reference point available, B_{pa}, for much longer than previously thought, and the spawning biomass might have fallen below the one million tonne level. Inclusion of the egg survey index as an absolute or relative measure of biomass does not seem to alter this conclusion (Figure 5.5). The use of all years of egg survey data also has a small effect on the biomass trajectory estimated by the assessment model.

The stock outlook obtained from the kriging-based SSB index should, at this point, be taken with caution. A number of issues on the present application of ICA to the NEA mackerel stock remain controversial. The choices of control and fitting parameters made by the working group (Appendix C) might not be completely adequate for the new estimates. For example, a fixed weight is assigned to the index over the biomass trajectory determined by VPA. A possible alternative is to weight the abundance index by the ratio of the variances of the standardised biomass and catch indices (Patterson & Melvin, 1996), for which the variance estimates provided here could be applied. However, it was the intention here to replicate closely the present-day assessment of this stock, and provide an indication of the impact of the application of the proposed estimator of egg production. Further exploration of possible changes to the assessment procedure under the light of the results presented here would undoubtedly be an interesting line of future work. A complementary analysis that could also help analysing the impact of uncertainty of stock assessment would be a Monte Carlo simulation of ICA-estimated SSB. Samples from the posterior distribution of egg survey-derived SSBs could be used as input to ICA to generate an approximate indication of the uncertainty around the SSB estimates introduced by the egg surveys. This is presently not possible due to limitations in the software implementation of ICA.

5.5.3 *Future research*

As fisheries management increasingly moves from the rigid establishment of control measures based on point estimates of parameter values and reference points, to a risk-based framework (Hilborn *et al.*, 2001; Charles, 2002), assessment methods should incorporate all available information on variability of the quantities of interest. Estimators of the kind

presented here provide the basic tool for risk-based and probability-aware appraisal and management of fisheries.

Another methodology for spatio-temporal estimation that could be of future use in situations similar to those encountered in ichthyoplankton surveys within the Bayesian paradigm is that of dynamic linear mixed models (Guo, 2003; Huerta *et al.*, 2004). Generalised linear mixed models (GLMMs) have been widely applied in recent years in a Bayesian setting to problem involving longitudinal data, repeated measurements at a number of spatial locations. As previously explained, the temporal component is explicitly modelled by means of autoregressive (ARMA, ARIMA) and similar time series methods. Although present development in this area concentrates on settings where a repeated number of samples are taken at fixed locations, such as meteorological data from weather stations, their use in fisheries problems should be further investigated. In particular, conditioning time series on both environmental factors and on observations obtained in the immediate vicinity, could provide dynamic methods that not only improve estimation, but provide guidance on sampling design. From a Bayesian perspective, GLMMs are well studied, and efficient MCMC samplers could be developed that take advantage of conditional probabilities and conjugate formulations to increase efficiency (Brooks, 1998). Real-time analysis of an ongoing egg survey could lead to adaptable and flexible survey designs, well suited to the changing dynamics of spawning by pelagic fish.

Another approach that could be of use in this problem is that of reversible jump MCMC (Green, 1995; Brooks *et al.*, 2003). Model uncertainty is explicitly incorporated with this methodology in the fitting procedure, by extending typical MCMC samplers to situations where the parameter vectors differ not only on their values but also on their dimensions. Variogram model selection, attempted in the present study by applying a number of fitting statistics, would instead be carried out within the MCMC procedure. Probabilities can then be computed for each alternative model structure given the available data. The extension of reversible-jump MCMC to variogram model selection could be a notable contribution to the introduction of Bayesian methods in the geostatistical field. A relatively general formulation could be probably developed covering commonly used variogram models, in both hierarchical and non-hierarchical settings.

An important question that needs further exploration is that of survey design and its interaction with the estimator being applied. Ideally, such an analysis should be based upon simulated datasets, covering a range of sampling scenarios. Sampling could perhaps be concentrated along the known area of higher abundance and rely on accumulated experience to extrapolate in time and space. A design more suited to the GAM and geostatistical estimators could also be investigated that covered more homogeneously the space-time cube. We have seen how limited we presently are when attempting to assess estimator performance by the problems at generating simulated datasets that incorporate the necessary characteristics. The use of GAM models specifically tailored to the task of

simulation could eventually provide an answer. Furthermore, present analyses incorporating data from all developmental stages (D. Beare, Marine Laboratory Aberdeen, *pers. com.*) might also help modelling egg distribution in space and time.

Such an exercise should eventually incorporate the question of the cost of sampling, and the decrease in the risks associated with managing the stock with more accurate or updated information about its status. Ichthyoplankton egg surveys are an expensive and complex exercise. A rough estimate for the NEA mackerel pelagic egg survey falls around the £2 million mark (D. Reid, MLA, *pers. com.*). This amount needs to be put into context of the estimated market value of the stock, which for the UK alone is estimated at around £81 million in 2003, according to the most recent statistics (DEFRA, 2003).

Mention can be finally made of the difficulties encountered when carrying out simulation-based studies of the kind presented here outside of a cohesive software framework. Having to deal with a number of pieces of software and programming languages forces the researcher to create multiple ad-hoc interfaces between the various parts of the analysis. These might range from filters for import and output of results in certain formats, to the modification of source code, if available, to generate results tailored to the needs of the next calculation in the process. Fisheries assessment and modelling commonly involve a range of models and procedures originating from various fields of statistics and mathematics. Accommodating these methods and software within an unifying framework would be of great help for analyses like the ones presented here. Efforts are presently underway to generate such a framework around the R statistical language (R Development Core Team, 2004). This language provides a powerful, flexible and extendable platform on which to converge most fisheries-related models and software. Simulations for assessment and management would then be made much easier as such platform provides the researcher with the possibility for mixing various analyses and methodologies under a single procedure.

Appendix A

Notation

Parameter	Definition
ϕ	range
ϕ_{xy}	range in the spatial dimensions (x,y)
ϕ_z	range in the temporal dimension (z)
c	sill
c_0	nugget
$\gamma(h)$	semi-variogram
h	distance (nm)
t	time (days)
κ	shape parameter (Matérn model)

Appendix B

Acronyms

Acronym	
ACFM	Advisory Committee on Fisheries Management
AEP	Annual Egg Production
AEPM	Annual Egg Production Method
BIC	Bayesian Information Criterion
DEPM	Daily Egg Production Method
DFRM	Daily Fecundity Reduction Method
DIC	Deviance Information Criterion
GAM	General Additive Model
ICA	Integrated Catch Anlaysis
ICES	International Councial for the Exploration of the Sea
NAO	North Atlantic Oscillation
NEA	North East Atlantic
SST	Sea Surface Temperature
VPA	Virtual Population Analysis
WGMHSA	Working Group on Mackerel, Horse mackerel, Sardine and Anchovy

References

Alheit, J. 1988. Reproductive-biology of Sprat (*Sprattus-sprattus*) - factors determining annual egg-production. *ICES Journal of Marine Science*, 44(2): 162–168.

Anon. 1999. Co-ordination of the development of an improved method of measuring volume filtered by high speed plankton samplers. EU AIR3 CT94 1911. Institut fur Meereskunde, Kiel.

Armstrong, M., Connolly, P., Nash, R., Pawson, M., Alesworth, E., Coulahan, P., Dickey-Collas, M., Milligan, S., O'Neill, M., Whitthames, P., Woolner, L. 2001. An application of the annual egg production method to estimate the spawning biomass of cod (*Gadus morhua* L.), plaice (*Pleuronectes platessa* L.) and sole (*Solea solea* L.) in the Irish Sea. *ICES Journal of Marine Science*, 58: 183–203.

Augustin, N., Borchers, D., Clarke, E., Buckland, S., Walsh, M. 1998. Spatiotemporal modelling for the annual egg production method of stock assessment using generalized additive models. *Canadian Journal of Fisheries and Aquatic Sciences*, 55: 2608–2621.

Barange, M., Hampton, I., Roel, B. 1999. Trends in the abundance and distribution of anchovy and sardine on the South African continental shelf in the 1990s, deduced from acoustic surveys. *South African Journal of Marine Science*, 21: 367–391.

Beare, D. J., Reid, D. G. 2002. Investigating spatio-temporal change in spawning activity by Atlantic mackerel between 1977 and 1998 using generalized additive models. *ICES Journal of Marine Science*, 59: 711–724.

Berger, J. 1985. *Statistical decision theory and Bayesian analysis.* Springer-Verlag, New York.

Bergh, M., Butterworth, D. 1987. Towards rational harvesting of the South African anchovy considering survey imprecision and recruitment variability. *South African Journal of Marine Science*, 5: 937–951.

Berke, O. 1998. On spatiotemporal prediction for on-line monitoring data. *Communications in Statistics-Theory and Methods*, 27(9): 2343–2369.

Best, N., Cowles, M., Vines, K. 1996. CODA: Convergence Diagnostics and Output Analysis software for Gibbs sampling output. Version 0.30, MRC Biostatistics Unit.

Beverton, R. J. H., Holt, S. J. 1957. *On the dynamics of exploited fish populations.* Fish & Fisheries Series. Chapman & Hall, London.

Bez, N. 2002. Global fish abundance estimation from regular sampling: the geostatistical transitive method. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 1921–1931.

Bez, N., Rivoirard, J. 2000. On the role of sea surface temperature on the spatial distribution opf early stages of mackerel using inertiograms. *ICES Journal of Marine Science*, 57(2): 383–392.

Bez, N., Rivoirard, J. 2001. Transitive geostatistics to characterise spatial aggregation with diffuse limits: an application on mackerel ichtyoplankton. *Fisheries Research*, 50(1-2): 41–58.

Bez, N., Rivoirard, J., Guiblin, P., Walsh, W. 1997. Covariogram and related tools for structural analysis of fish survey data. *In* E.Y. Baafi. N.A. Schofield (eds.), *Geostatistics Wollongong '96*, volume 2, pp. 1316–1327.

Bogaert, P. 1996. Comparison of kriging techniques in a space-time context. *Mathematical Geology*, 28: 73–86.

Booth, A. 2000. Incorporating the spatial component of fisheries data into stock assessment models. *ICES Journal of Marine Science*, 57(4): 858–865.

Borja, A., Uriarte, A., Egana, J. 2002. Environmental factors and recruitment of mackerel, Scomber scombrus L. 1758, along the north-east Atlantic coasts of Europe. *Fisheries Oceanography*, 11: 116–127.

Box, G. E. P., Cox, D. R. 1964. An analysis of transformations. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 26: 211–246.

Box, G. E. P., Tiao, G. C. 1973. *Bayesian inference in statistical analysis*. John Wiley & Sons, New York.

Brander, K. M., Milligan, S. P., Nichols, J. H. 1993. Flume Tank Experiments to Estimate the Volume Filtered By High- Speed Plankton Samplers and to Assess the Effect of Net Clogging. *Journal of Plankton Research*, 15(4): 385–401.

Breen, P. A., Kim, S. W., Andrew, N. L. 2003. A length-based Bayesian stock assessment model for the New Zealand abalone Haliotis iris. *Mar. Freshw. Res.*, 54: 619–634.

Brooks, S. P. 1998. Markov chain Monte Carlo method and its application. *Journal of the Royal Statistical Society Series D-The Statistician*, 47: 69–100.

Brooks, S. P., Giudici, P., Roberts, G. O. 2003. Efficient construction of reversible jump Markov chain Monte Carlo proposal distributions. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 65: 3–39.

Brown, P., Le, N., Zidek, J. 1994. Multivariate spatial interpolation and exposure to air pollutants. *Canadian Journal of Statistics*, 26: 211–252.

Bull, B., Francis, R., Dunn, A., McKenzie, A., Gilbert, D., Smith, M. 2005. CASAL (C++ algorithmic stock assessment laboratory): CASAL user manual v2.07-2005/08/21. Technical Report 127, NIWA.

Bulman, C., Koslow, J., Haskard, K. 1999. Estimation of the spawning stock biomass of blue grenadier (Macruronus novaezelandiae) off western Tasmania based upon the annual egg production method. *Marine and Freshwater Research*, 50(3): 197–207.

Carr, J., Palmer, J. 1993. Revisiting the accurate calculation of block-sample covariances using gauss quadrature. *Mathematical Geology*, 25(5): 507–524.

Charles, A. T. 2002. The precautionary approach and 'burden of proof' challenges in fishery management. *Bulletin of Marine Science*, 70: 683–694.

Chaudhuri, A., Fuentes, M., Holland, D. 2004. Predicting nonstationary processes using Reversible MCMC. Internal Mimeo Series 2553, Department of Statistics, University of North Carolina.

Christakos, G. 2000. *Modern Spatiotemporal Geostatistics.* OUP.

Christakos, G., Li, X. 1998. Bayesian maximum entropy analysis and mapping: A farewell to kriging estimators? *Mathematical Geology*, 30: 435–462.

Conan, G. 1985. Assessment of shellfish stocks by geostatistical techniques. CM 1984/K:30, ICES.

Congdon, P. 2001. *Bayesian statistical modelling.* John Wiley & Sons.

Cooke, J. G. 1999. Improvement of fishery-management advice through simulation testing of harvest algorithms. *ICES Journal of Marine Science*, 56: 797–810.

Coombs, S. H., Morgans, D., Halliday, N. C. 2001. Seasonal and ontogenetic changes in the vertical distribution of eggs and larvae of mackerel (Scomber scombrus L.) and horse mackerel (Trachurus trachurus L.). *Fisheries Research*, 50(1-2): 27–40.

Cornish, T. 1883. Mackerel and pilchard fisheries. *In Papers of the conferences held in connection with the great International Fisheries Exhibition.* W. Clowes and Sons, London.

Cressie, N., Hawkins, D. 1980. Robust estimation of the variogram. *Journal of the International Association for Mathematical Geology*, 12: 115–125.

Cressie, N. A. 1991. *Statistics for spatial data.* Wiley & Sons, New York.

Curtis, K. A. 2004. Fine scale spatial pattern of pacific sardine (*Sardinops sagax*) and northern anchovy (*Engraulis mordax*) eggs. *Fisheries Oceanography*, 13(4): 239–239.

De Oliveira, V., Kedem, B., Short, D. A. 1997. Bayesian prediction of transformed Gaussian random fields. *Journal of the American Statistical Association*, 92: 1422–1433.

DEFRA. 2003. United Kingdom Sea Fisheries Statistics. Available at http://statistics.defra.gov.uk/esg/publications/fishstat/.

Deutsch, C. V., Journel, A. G. 1998. *GSLIB: Geostatistical software library and user's guide.* Oxford University Press, 2nd edition.

Diggle, P., Tawn, J., Moyeed, R. 1998. Model-based geostatistics. *Journal of the Royal Statistical Society Series C-Applied Statistics*, 47: 299–326.

Eaton, D. 1989. Spawning-stock biomass of scad (Trachurus trachurus L.) to the west of the British Isles, as indicated by egg surveys. *Journal du Conseil ICES*, 45(3): 231–247.

Emerson, L., Greer Walker, M., Witthames, P. 1990. A stereological method for estimating fish fecundity. *Journal of Fish Biology*, 36: 721–730.

FAO. 1999. Guidelines for the Routine Collection of Capture Fishery Data. FAO Fisheries Technical Paper 382, Food & Agricultural Organization of the United Nations, Roma.

Fariña, A. C., Freire, J., Gurriarán, E. G. 1994. *Nephrops norvegicus* in the Galician continental shelf (NW Spain): abundance and distribution. *Fisheries Research*, 19(3-4): 333–347.

Findlay, M., Lansley, J. 1998. Plankton sample colour as a determinant of the abundance of stage I eggs in the western mackerel (*Scomber scombrus* L.) stock. *ICES Journal of Marine Science*, 55: 141–144.

Fletcher, W. J., Sumner, N. R. 1999. Spatial distribution of sardine (*Sardinops sagax*) eggs and larvae: an application of geostatistics and resampling to survey data. *Canadian Journal of Fisheries and Aquatic Sciences*, 56(6): 907–914.

Fox, C. J., O'Brien, C. M., Dickey-Collas, M., Nash, R. D. M. 2000. Patterns in the spawning of cod (Gadus morhua L.), sole (Solea solea L.) and plaice (Pleuronectes platessa L.) in the Irish Sea as determined by generalized additive modelling. *Fisheries Oceanography*, 9: 33–49.

Freire, J., González Gurriáran, E., Olaso, I. 1992. Spatial distribution of *Munida intermedia* and *M. sars*, (Crustacea: Anomura) on the Galician continental shelf (NW Spain): application of geostatistical analysis. *Estuarine, Coastal & Shelf Science*. 35(6): 637–648.

Freire, J., González Gurriarán, E., Olaso, I., Fernández, L. 1991. Análisis geostadístico de la distribución espacial de *Plesionika heterocarpus* y *Solenocera membranacea* (Crustacea: Decapoda) en la plataforma continental de Galicia (NW España). *Boletín del Instituto Español de Oceanografía*, 7(2): 79–88.

Freon, P., Misund, O. A. 1999. *Dynamics of pelagic fish distribution and behaviour: Effects on fisheries and stock assessment.* Fishing News Books, Oxford (UK).

Gamerman, D. 1997. *Markov Chain Monte Carlo: Stochastic simulation for Bayesian inference.* Chapman & Hall, London.

Gaudard, M., Karson, M., Linder, E., Sinha, D. 1999. Bayesian spatial prediction. *Environmental & Ecological Statistics*, 6(2): 147–171.

Gelman, A., Carlin, J. B., Stern, H. S., Rubin, D. B. 1995. *Bayesian data analysis.* Texts in Statistical Sciences. Chapman & Hall, London.

Gelman, A., Rubin, D. 1992. Inference from iterative simulation using multiple sequences. *Statistical Science*, 7: 457–511.

Geweke, J. 1992. Evaluating the accuracy of sampling-based approaches to calculating posterior moments. *In* J. Bernardo, J. Berger, A. Dawid, A. Smith (eds.), *Bayesian Statistics 4*. Clarendon Press, Oxford, UK.

Gohin, F. 1985. Planification des expériencies et interpritation par la théorie des variables régionaliseés: application á l'estimation de la biomasse d'una plage. CM 1985/D:3. ICES.

Goovaerts, P. 1997. *Geostatistics for natural resources evaluation.* OUP, New York.

Green, P. J. 1995. Reversible jump Markov chain Monte Carlo computation and Bayesian model determination. *Biometrics*, 82(4): 711–732.

Grégoire, F., Lambert, J., Lévesque, C., Beaulieu, M. F. 1997. Assessment of the Atlantic mackerel (*Scomber scombrus*), biomass for 1996 by the Total Eggs Production Method. Research Document 97/71, Department of Fisheries and Oceans. Canadian Stock Assessment Secretariat.

Guillard, J., Gerdeaux, D., Brun, G., Chappaz, R. 1992. The use of geostatistics to analyse data from an echo-integration survey of fish stock in Lake Saint-Croix. *Fisheries Research*, 13: 395–406.

Gunderson, D. 1993. *Surveys of fisheries resources.* John Wiley & Sons, New York.

Guo, W. S. 2003. Dynamic state-space models. *Journal of Time Series Analysis*, 24: 149–158.

Gurriarán, E. G. 1993. Geostatistical analysis of spatial distribution of *Liocarcinus depurator*, *Macropipus tuberculatus*, and *Polybius henslowii* (Crustacea, Brachyura) over the Galician continental shelf (NW Spain). *Marine Biology*, 115(3): 453–461.

Handcock, M., Stein, M. 1993. A Bayesian-Analysis of Kriging. *Technometrics*, 35(4): 403–410.

Handcock, M., Wallis, J. 1994. An approach to statistical spatio-temporal modeling of meteorological fields. *Journal of the American Statistical Association*, 89: 368–378.

Harley, S. J., Myers, R. A. 2001. Hierarchical Bayesian models of length-specific catchability of research trawl surveys. *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 1569–1584.

Hart, P. J. B., Reynolds, J. D. 2002. *The Handbook of Fish Biology and Fisheries.* Blackwell, Oxford.

Hastings, W. 1970. Monte Carlo sampling methods using Markov chains and their applications. *Biometrika*, 57(97-109).

Heidelberger, P., Welch, P. 1983. Simulation run length control in the presence of an initial transient. *Operations Research*, 31: 1109–1144.

Hensen, V., Apstein, C. 1897. *Die Nordsee-Expedition 1895 des Deutschen Seefischereivereins. Über die Eiemenge der im Winter Laichenden Fische.* Wiss. Meeresunter., Helgoland.

Hilborn, R. 2003. The state of the art in stock assessment: where we are and where we are going. *Scientia Marina*, 67: 15–20.

Hilborn, R., Maguire, J.-J., Parma, A. 2001. The precautonary approach and risk management: can they increase the probability of successes in fishery management? *Canadian Journal of Fisheries and Aquatic Sciences*, 58(1): 99–107.

Hilborn, R., Walters, C. J. 1992. *Quantitative fisheries stock assessment : choice, dynamics, and uncertainty.* Chapman & Hall, New York.

Hoenig, J. M., Warren, W. G., Stocker, M. 1994. Bayesian and related approaches to fitting surplus production models. *Canadian Journal of Fisheries and Aquatic Sciences*, 51: 1823–1831.

Huerta, G., Sansó, B., Stroud, J. 2004. A spatiotemporal model for Mexico City ozone levels. *Journal of the Royal Statistical Society Series C-Applied Statistics*, 53(2): 231–248.

Hunter, J., Lo, N.-H. 1993. Ichthyoplankton methods for estimating fish biomass introduction and terminology. *Bulletin of Marine Science*, 53(2): 723–727.

ICES. 1979. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 1979/H:5, ICES, Copenhagen.

ICES. 1981. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 1981/H:7, ICES, Copenhagen.

ICES. 1984. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 1984/H:3, ICES, Copenhagen.

ICES. 1987. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 1997/H:2, ICES, Copenhagen.

ICES. 1990. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 1990/H:2, ICES, Copenhagen.

ICES. 1996. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 1996/H:8, ICES, Copenhagen.

ICES. 1999a. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 1999/G:5, ICES.

ICES. 1999b. Report of the Working Group on the Assessment of Mackerel. Horse Mackerel, Sardine and Anchovy. CM 1999/ACFM:6, ICES, Copenhagen.

ICES. 2000. Report of the Working Group on Mackerel and Horse Mackerel Egg Surveys. CM 2000/G:01, ICES, Copenhagen.

ICES. 2001. Mackerel and Horse Mackerel Egg Staging and Histology Workshop. CM 2001/G:01, ICES.

ICES. 2004a. Advisory Committee on Fisheries Management Annual Report. *ICES Advice*. Vol. 1, No. 2, ICES, Copenhagen, Denmark.

ICES. 2004b. Report of the Working Group on the Assessment of Mackerel, Horse Mackerel, Sardine and Anchovy. CM 2004/ACFM:8, ICES.

Imrie, C., Korre, A., Reid, D. 2003. Estimation of the total annual egg production and associated uncertainty of Atlantic mackerel (*Scomber scombrus*) using conditional simulation. Working Document. Meeting of the Working Group on Mackerel and Horse Mackerel Egg Survey. Lisbon, 1-4 April, ICES.

Isaaks, E. H., Srivastiva, R. M. 1989. *An introduction to applied geostatistics.* Oxford University Press.

Journel, A., Huijbregts, C. 1978. *Mining geostatistics.* Academic Press, London.

Kim, S., Gunderson, D. 1989. Cohort dynamics of walleye pollock in Shelikof Strait, Gulf of Alaska, during the egg and larval periods. *Transactions of the American Fisheries Society*, 118(3): 264–273.

King, R., Brooks, S. P. 2002. Bayesian model discrimination for multiple strata capturerecapture data. *Biometrika*, 89: 785–806.

Kiridiakis, P., Journel, A. 1999. Geostatistical space-time models: a review. *Mathematical Geology*, 31(6): 651–684.

Kitanidis, P. K. 1997. *Introduction to geostatistics: Applications to hydrogeology.* Cambridge University Press.

Kitanidis, P. K., Shen, K. F. 1996. Geostatistical interpolation of chemical concentration. *Advances in Water Resources*, 19: 369–378.

Krebs, C. 1989. *Ecological methodology.* Harper & Row, New York.

Kyriakidis, P. C., Journel, A. G. 1999. Geostatistical space-time models: A review. *Mathematical Geology*, 31: 651–684.

Lee, P. M. 1997. *Bayesian statistics: An introduction.* Arnold, London, 2nd edition.

Lockwood, S. 1988. *The mackerel : its biology, assessment and the management of a fishery.* Fishing News Books, Oxford.

Lockwood, S., Nichols, J., Coombs, S. 1977. The development rates of mackerel (Scomber scombrus, L.) eggs over a range of temperature. CM 1977/J:13, ICES.

Lockwood, S., Nichols, J., Dawson, W. 1981. The estimation of mackerel (*Scomber scombrus* L.) spawning stock size by plankton survey. *Journal of Plankton Research*, 3(2): 217–233.

MacCall, A. 2001. Sardine regimes and mesoscale structure (an integrative hypothesis). *In* A. Bakun, K. Broad (eds.), *The IRI-IPRC Pacific Climate-Fisheries Workshop. Honolulu*, IRI CW/02/1, pp. 39–42. Palisades, NY.

Maravelias, C. D., Reid, D. G., Simmonds, E. J., Haralabous, J. 1996. Spatial analysis and mapping of accoustic survey data in the presence of high local variability: geostatistical application to North Sea herring (*Clupea harengus*). *Canadian Journal of Fisheries and Aquatic Sciences*, 53: 1497–1505.

Mátern, B. 1986. *Spatial variation*, *Lecture Notes in Statistics*, volume 36. Springer-Verlag., New York, 2nd edition.

Maynou, F. X., Sardá, F. 1998. Assessment of the spatial structure and biomass evaluation of *Nephrops norvegicus* populations in the northwestern Mediterranean by geostatistics. *ICES Journal of Marine Science*, 55: 102–120.

McAllister, M., Ianelli, J. 1997. Bayesian stock assessment using catch-age data and the sampling - Importance resampling algorithm. *Canadian Journal of Fisheries and Aquatic Sciences*, 54(2): 284–300.

McAllister, M., Pikitch, E., Punt, A., Hilborn, R. 1994. A Bayesian-Approach to Stock Assessment and Harvest Decisions Using the Sampling Importance Resampling Algorithm. *Canadian Journal of Fisheries and Aquatic Sciences*, 51(12): 2673–2687.

McAllister, M. K., Hill, S. L., Agnew, D. J., Kirkwood, G. P., Beddington, J. R. 2004. A Bayesian hierarchical formulation of the De Lury stock assessment model for abundance estimation of Falkland Islands' squid (Loligo gahi). *Canadian Journal of Fisheries and Aquatic Sciences*, 61(6): 1048–1059.

McCullagh, P., Nelder, J. 1989. *Generalized linear models*. Chapman & Hall, London, 2nd edition.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E. 1953. Equations of State Calculations by Fast Computing Machines. *Journal of Chemical Physics*, 21: 1087–1091.

Meyer, R., Millar, R. B. 1999. Bayesian stock assessment using a state-space implementation of the delay difference model. *Canadian Journal of Fisheries and Aquatic Sciences*, 56: 37–52.

Michielsens, C. G., McAllister, M. K. 2004. A Bayesian hierarchical analysis of stock-recruit data: quantifying structural and parameter uncertainties. *Canadian Journal of Fisheries and Aquatic Sciences*, 61(6): 1032–1047.

Millar, R. B., Meyer, R. 2000. Non-linear state space modelling of fisheries biomass dynamics by using Metropolis-Hastings within-Gibbs sampling. *Journal of the Royal Statistical Society Series C-Applied Statistics*, 49: 327–342.

Ming Ye, Shlomo P. Neuman, Philip D. Meyer. 2004. Maximum likelihood Bayesian averaging of spatial variability models in unsaturated fractured tuff. *Water Resources Research*, 40: 23–36.

Myers, R. A., Mertz, G. 1998. Reducing uncertainty in the biological basis of fisheries management by meta-analysis of data from many populations: a synthesis. *Fisheries Research*, 37: 51–60.

Omre, H. 1987. Bayesian Kriging-Merging Observations and Qualified Guesses in Kriging. *Mathematical Geology*, 19(1): 25–39.

Omre, H., Halvorsen, K. 1989. The Bayesian Bridge Between Simple and Universal Kriging. *Mathematical Geology*, 21(7): 767–786.

Pardo Iguzquiza, E. 1999. Bayesian inference of spatial covariance parameters. *Mathematical Geology*, 31(1): 47–65.

Patterson, K. R., Melvin, G. 1996. Integrated Catch at Age Analysis Version 1.4. Scottish Fisheries Research Report, No 58, Scottish Fisheries Research Services.

Pauly, D., Christensen, V., Dalsgaard, J., Froese, R., Torres, F. 1998. Fishing down marine food webs. *Science*, 279: 860–863.

Pebesma, E. J., Wesseling, C. G. 1998. Gstat: A program for geostatistical modelling, prediction and simulation. *Computers and Geosciences*, 24: 17–31.

Pelletier, D., Parma, A. M. 1994. Spatial-distribution of pacific halibut (*Hippoglossus stenolepis*) - an application of geostatistics to longline survey data. *Canadian Journal of Fisheries and Aquatic Sciences*, 51: 1506–1518.

Petitgas, P. 1993. Geostatistics for fish stock assessments - a review and an acoustic application. *ICES Journal of Marine Science*, 50(3): 285–298.

Petitgas, P. 1997. Sole egg distributions in space and time characterised by a geostatistical model and its estimation variance. *ICES Journal of Marine Science*, 54(2): 213–225.

Petitgas, P. 2001. Geostatistics in fisheries survey design and stock assessment: models, variances and applications. *Fish and Fisheries*, 2: 231–249.

Petitgas, P., Poulard, J. 1989. Applying stationary geostatistics to fisheries: A study on hake in the Bay of Biscay. CM 1989:S 8, ICES.

Picquelle, S., Megrey, B. 1993. A preliminary spawning biomass estimate of walleye pollock, Theragra chalcogramma, in the Shelikof Strait, Alaska, based on the annual egg production method. *Bulletin of Marine Science*, 53(2): 728–749.

Pipe, R., Walker, P. 1987. The effect of temperature on the development of Scad (*Trachurus trachurus* L.) eggs. *Journal of Fish Biology*, 31: 675–682.

Punt, A., Hilborn, R. 1997. Fisheries stock assessment and decision analysis: The Bayesian approach. *Review in Fish Biology & Fisheries*, 7(1): 35–63.

Qian, S. 1997. Estimating the area affected by phosphorus runoff in an Everglades wetland: A comparison of universal kriging and Bayesian kriging. *Environmental and Ecological Statistics*, 4(1): 1–26.

Quinn, T. J., Deriso, R. B. 1999. *Quantitative fish dynamics*. Oxford University Press.

R Development Core Team. 2004. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna.

Raftery, A., Lewis, S. 1992. One long run with diagnostics: Implementation strategies for Markov chain Monte Carlo. *Statistical Science*, 7: 493–497.

Reid, D. G., Walsh, M., Turrell, W. R. 2001. Hydrography and mackerel distribution on the shelf edge west of the Norwegian deeps. *Fisheries Research*, 50(1-2): 141–150.

Restrepo, V. 1995. Application of cohort slicing and tuned VPA to simulated data that includes variability in length at age. *Collective Volume of Scientific Papers ICCAT.* 44(2): 67–71.

Restrepo, V., Powers, J., Turner, S. 1990. Incorporating uncertainty in VPA results via simulation. *Collective Volume of Scientific Papers ICCAT*, 35(2): 355–361.

Ribeiro, P. J., Diggle, P. J. 2000. Bayesian inference in Gaussian model-based geostatistics. ST-99-08, Department of Mathematics and Statistics, Lancaster University.

Rivoirard, J., Simmonds, E., Foote, K., Fernandez, P., Bez, N. 2000. *Geostatistics for estimating fish abundance.* Blackwell Science, Oxford.

Saville, A. 1964. Estimation of the abundance of a fish stock from egg and larval surveys. Rapport V Runion, 155 (164-173). Conseil International pour l'Exploration de la Mer.

Schwarz, G. 1978. Estimating the dimension of a model. *The Annals of Statistics.* 6(2): 462–464.

Shelton, P., Armstrong, M., Roel, B. 1993. An overview of the application of the daily egg production method in the assessment and management of anchovy in the Southeast Atlantic. *Bulletin of Marine Science*, 53(21): 778–794.

Simmonds, E., Beare, D., Reid, D. 2003. Sensitivity of the current ICA assessment of western mackerel and short term prediction to the sampling error in the egg survey parameters. CM 2003/X:10, International Council for the Exploration of the Sea.

Slotte, A. 2004. Historic changes in the condition of NEA mackerel: Possible effects on fecundity. Working document WGMHAS 21/03, ICES.

Smith, A., Bernardo, J. 2000. *Bayesian theory.* John Wiley and Sons, Chichester.

Spencer, P. D., Collie, J. S. 1997. Patterns of population variability in marine fish stocks. *Fisheries Oceanography*, 6: 188–204.

Spiegelhalter, D., Best, N., Carlin, B., van der Linde, A. 2002. Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 64(4): 583–639.

Stein, M. 1999. *Interpolation of spatial data: Some theory for kriging.* Springer series in statistics. Springer-Verlag, New York.

Steinhaus, H. 1999. *Mathematical Snapshots.* Dover, New York, 3rd edition.

Stroud, J. R., Müller, P., Sansó, B. 2001. Dynamic models for spatiotemporal data. *Journal of the Royal Statistical Society Series B-Statistical Methodology*, 63: 673–698.

Sullivan, P. 1991. Stock abundance estimation using depth-dependent trends and spatially correlated variation. *Canadian Journal of Fisheries and Aquatic Sciences*, 48(9): 1691–1703.

Sundby, S., Solemdal, P. 1984. The egg production of Arcto-Norwegian cod (Gadus morhua L.) in the Lofoten area estimated by egg surveys. *In* O. Godoe, S. Tilseth (eds.), *Proceeding of the Soviet-Norwegian Symposium on Reproduction and Recruitment of Arctic Cod*, pp. 113–135. Institute of Marine Research, Institute of Marine Research, Bergen (Norway).

Todini, E. 2001. A Bayesian technique for conditioning radar precipitation estimates to rain-gauge measurements. *Hydrology & Earth System Sciences.* 5: 187–199.

Uriarte, A., Lucio, P. 2001. Migration of adult mackerel along the Atlantic European shelf edge from a tagging experiment in the south of the Bay of Biscay in 1994. *Fisheries Research*, 50(1-2): 129–139.

Walters, C., Hilborn, R. 1976. Adaptive control of fishing systems. *J. Fish. Res. Bd. Can.*, 33: 145–59.

Walters, C. J., Martell, S. J. D. 2004. *Fisheries Ecology and Management.* Princeton UP.

Witthames, P. R. 2000. *A manual for the estimation of fecundity and atresia in Mackerel and Horse mackerel.* CEFAS, Lowestoft.

Zar, J. H. 1996. *Biostatistical analysis.* Prentice-Hall, 3rd edition.