

Setting biological reference points for Atlantic salmon stocks: transfer of information from data-rich to sparse-data situations by Bayesian hierarchical modelling

Etienne Prévost, Eric Parent, Walter Crozier, Ian Davidson, Jacques Dumas, Gudni Gudbergsson, Kjetil Hindar, Phil McGinnity, Julian MacLean, and Leif M. Sættem

Prévost, E., Parent, E., Crozier, W., Davidson, I., Dumas, J., Gudbergsson, G., Hindar, K., McGinnity, P., MacLean, J., and Sættem, L. M. 2003. Setting biological reference points for Atlantic salmon stocks: transfer of information from data-rich to sparse-data situations by Bayesian hierarchical modelling. — ICES Journal of Marine Science, 60: 1177–1193.

We present an application of Bayesian hierarchical modelling of stock–recruitment (SR) relationships aiming at estimating Biological Reference Points (BRP) for European Atlantic salmon (*Salmo salar*) stocks. The structure of the hierarchical SR model developed distinguishes two nested levels of randomness, within-river and between rivers. It is an extension of the classical Ricker model, where the parameters of the Ricker function are assumed to be different between rivers, but drawn from a common probability distribution conditionally on two covariates: river size and latitude. The output of ultimate interest is the posterior predictive distribution of the SR parameters and their associated BRP for a new river with no SR data.

The flexible framework of the Bayesian hierarchical SR analysis is a step towards making the most comprehensive use of detailed stock monitoring programs for improving management advice. Posterior predictive inferences may be imprecise due to the relative paucity of information introduced in the analysis compared to the variability of the stochastic process modeled. Even in such cases, direct extrapolation of results from local data-rich stocks should be dismissed as it can lead to a major underestimation of our uncertainty about management parameters in sparse-data situations. The aggregation of several stocks under a regional complex improves the precision of the posterior predictive inferences. When several stocks are managed jointly, even imprecise knowledge about each component of the aggregate can be valuable. The introduction of covariates to explain between stock variations provides a significant gain in the precision of the posterior predictive inferences. Because we must be able to measure the covariates for all the stocks of interest, i.e. mostly sparse-data cases, the number of covariates which can be used in practice is limited. The definition of the assemblage of stocks which we model as exchangeable units, conditionally on the covariates, remains the most influential choice to be made when attempting to transfer information from data-rich to sparse-data situations.

© 2003 International Council for the Exploration of the Sea. Published by Elsevier Ltd. All rights reserved.

Keywords: Atlantic salmon, Bayesian, biological reference points, hierarchical model, stock and recruitment.

Received 24 March 2003; accepted 27 August 2003.

E. Prévost: INRA, UMR EQHC, 65 rue de St Brieuc, 35042 Rennes cedex, France.
E. Parent: GRESE laboratory, ENGREF, 19 avenue du Maine, 75732 Paris cedex 15, France.
W. Crozier: Agricultural and Environmental Sciences Division, DARDNI, Newforge Lane, Belfast BT9 5PX, Northern Ireland, UK. I. Davidson: Environment Agency, National Salmon and Trout Fisheries Centre, Rivers House, St Mellons Business Park, Cardiff CF3 0LT, Wales, UK. J. Dumas: INRA, UMR ECOBIOP, BP 3, 64310 St Pée s/Nivelle, France.
G. Gudbergsson: Institute of Freshwater Fisheries, Vagnhöfða, Reykjavík 112, Iceland.
K. Hindar: Norwegian Institute for Nature Research, Tungasletta 2, 7485 Trondheim, Norway. P. McGinnity: Salmon Research Agency of Ireland, Furnace, Newport, Co. Mayo, Ireland. J. MacLean: Freshwater Fisheries Laboratory, Field Station, 16 River Street, Montrose, Angus DD10 8DL, Scotland, UK. L. M. Sættem: Department of Environment, County Governor of Møre and Romsdal, 6404 Molde, Norway. Correspondence to E. Prévost; tel: +33 2 23 48 52 48; fax: +33 2 23 48 54 40; e-mail: prevost@roazhon.inra.fr.

Introduction

Over the recent years, the Bayesian treatment of hierarchical models has gained increasing interest in fisheries science (Hilborn and Liermann, 1998). It has been applied to a wide range of problems, mainly population size estimation (Su *et al.*, 2001; Rivot and Prévost, 2002; Wyatt, 2002), meta-analysis of fundamental population dynamics processes (Liermann and Hilborn, 1997; Myers, 2001) and Biological Reference Points estimation (Myers and Mertz, 1998; Prévost *et al.*, 2001; Chen and Holtby, 2002; Dorn, 2002). Covariates can be introduced into hierarchical models to explain variations between stocks and improve the estimation precision of quantities of interest (Myers *et al.*, 2001, 2002). We present here an additional application of Bayesian hierarchical modelling aiming at estimating Biological Reference Points for European Atlantic salmon (*Salmo salar*) stocks.

The analysis of stock and recruitment (SR) data is the most widely used approach for deriving Biological Reference Points for salmon (Prévost and Chaput, 2001). There are several hundreds of salmon stocks in the north east Atlantic area, each having its own characteristics with regard to the SR relationship. But resources to collect SR data are limited and there are only a restricted number of monitored rivers. In these rivers, adult returns, spawning escapement and sometimes smolt production are estimated yearly. Suitable SR series (both in terms of length and reliability of observations) are available for only a handful of monitored rivers spread throughout the European area of distribution of the species. Extrapolation of knowledge gained from monitored rivers to rivers for which SR data are not available is therefore required.

How can the SR information from the monitored (i.e. data-rich) rivers be used to set Biological Reference Points for other (i.e. sparse-data) salmon rivers while accounting for the major sources of uncertainty? Until recently, this question was essentially addressed in practice by extrapolating the Biological Reference Points determined from a single river SR series to an entire region or country while accounting only for the variations of size between rivers (see the review of Prévost *et al.*, 2001). When SR data are available from several rivers which are considered to be representative of an assemblage of rivers, the above question can be reformulated as follows: what can be inferred about the nature of the SR relationship for any new river of the assemblage based on the data from the sampled rivers? There are two nested sources of uncertainty in this situation. The first level of uncertainty is associated with the fact that there is relevant SR information available from a limited number of rivers within the assemblage of rivers. The second level of uncertainty relates to the limited number of SR observations available within each river. The Bayesian treatment of a hierarchical SR model, which we term a Bayesian hierarchical SR analysis, provides a rigorous framework for integrating these two nested levels of un-

certainty to derive a probability distribution of Biological Reference Points for a river with no SR data.

Walters and Korman (2001) give a full and critical exposure of the procedures relying on the classical SR models used for analyzing a single stock data set. They recommend the use of simple models for describing the stochastic relation between the spawning stock and the subsequent recruitment. They also advocate the adoption of the Bayesian approach for uncertainty assessment: the knowledge/uncertainty about Biological Reference Points should be reflected by probability distributions given the SR data in hand. From this basis, we propose an extension of the classical single stock SR models allowing the joint treatment of several SR data sets. Our hierarchical model is similar to those already presented by some authors (e.g. Myers *et al.*, 2001, 2002; Dorn, 2002). At the within-river level, it relies on a stochastic SR model depending on few parameters. At the between-river level, variations in SR parameters are again modeled according to a stochastic model conditioned by environmental covariates. After having assigned little informative prior probability distributions, we proceed to the Bayesian treatment of our hierarchical model using a set of 13 SR series gathered from monitored salmon rivers located in western Europe. Contrary to many studies which relied on partially Bayesian approaches (Liermann and Hilborn, 1997; Myers and Mertz, 1998; Myers, 2001; Chen and Holtby, 2002), we adopt a fully Bayesian setting for a more complete accounting of uncertainty. Thanks to Monte Carlo Markov Chain (MCMC) sampling techniques (Link *et al.*, 2002) currently available for the estimation of posterior distributions, a fully Bayesian approach can be implemented without complicating the practical realization of the analysis.

Material and methods

Rationale and fundamentals of the Bayesian hierarchical SR analysis

The Bayesian approach provides a consistent framework which allows the formulation of direct probabilistic statements about various sorts of unknowns, e.g. model parameters, missing data, unobservable variables; it is the method of choice to deal with uncertainty (Berger, 1985; Gelman *et al.*, 1995).

To make inferences from data in a Bayesian framework, a probabilistic (i.e. stochastic) model, representing our prior understanding of the process generating the observed data, must be set. Here, our data are SR observations. We can use standard SR models such as a Ricker curve with lognormal random errors to represent the link between the stock and the subsequent recruitment within any single river. Such a single river SR model is controlled by a few parameters, which directly are Biological Reference Points or from which Biological Reference Points can be computed (Walters and Korman, 2001). We denote by θ_i the SR

parameters vector of the river i . In our case, inference based on the data from the data-rich rivers about any sparse-data river is of special interest. The model must therefore specify the link between salmon rivers irrespective of whether SR data are available for them. We translate the idea that all salmon rivers belong to a common assemblage of rivers by considering the θ_i 's as realizations from a common probability distribution. This probability distribution is itself controlled by parameters, also called hyper-parameters. We denote by Θ the vector of hyper-parameters.

The conditioning structure corresponding to this general setting can be represented by a Directed Acyclic Graph (Figure 1; Spiegelhalter *et al.*, 1996). It is a hierarchical setting because:

- the distribution of the recruitment for any given level of stock is controlled by the θ_i 's parameters, and
- the distribution of the θ_i 's parameters is controlled by the Θ hyper-parameters.

This hierarchical structure organizes the transfer of information sourced by the data-rich rivers SR data towards the sparse-data rivers. The SR data from the monitored rivers improve our knowledge about the θ_i 's. This information gained about the θ_i 's allows in turn to improve our knowledge about Θ . The information gained on Θ provides insight into the SR parameters of any new river for which no SR data are available.

The hierarchical setting is midway between a complete pooling of SR data sets and the independent treatment of each single river SR series. Complete pooling of SR data sets relies on the assumption that there is a unique SR relationship common to all rivers, i.e. $\theta_i = \theta_j$ for any $i \neq j$. This is certainly an oversimplifying assumption. Conversely, full independence between rivers would mean that there is nothing to learn from the monitored rivers about the SR relationship of the other rivers. This is not sensible either and contradictory to the very essence of monitored river projects which aim at gaining knowledge useful in a wider context. By considering the θ_i 's as realizations from a common probability distribution, we acknowledge they can be different between rivers while at the same time related. This intermediate assumption opens the door for learning through the transfer of information between rivers. Any gain of information about a θ_i consequentially provides information about the probability distribution of the θ_i 's, thus bringing information about any θ_j , $j \neq i$. The Bayesian treatment of a hierarchical model allows the data to tell us how much can be learned from the monitored rivers.

Implicit but crucial in what has just been introduced is the hypothesis of exchangeability of the rivers with regard to their SR parameters. In our case, it means that, apart from the SR data, we have no insight into the phenomena causing variations in the SR relationship among rivers. In

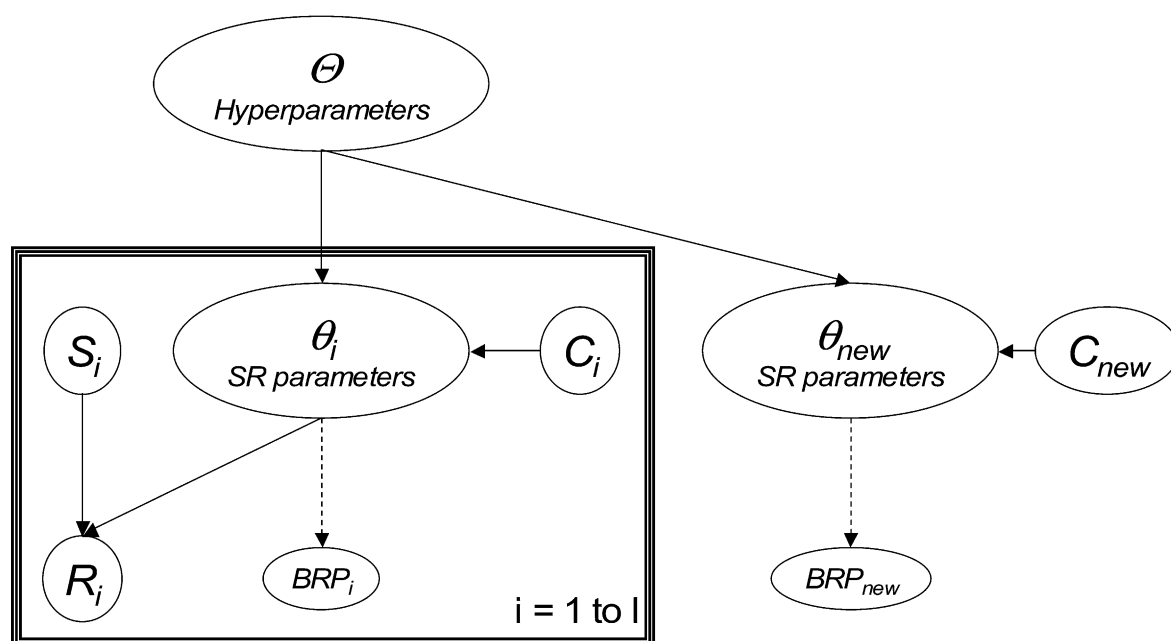


Figure 1. The conditioning structure of the Bayesian hierarchical SR analysis as represented in a Directed Acyclic Graph. Nodes (ellipses) are random variables. The plain arrows represent stochastic links, i.e. the distribution of a child node depends on its parents. Dashed arrows represent deterministic links, i.e. the Biological Reference Points are functions of the θ_i 's. S_i and R_i are the series of observed stock and recruitment for the monitored river i . C_i is a vector of explanatory covariate of the θ_i 's. The frame means there are I monitored rivers with SR data. The "new" index refers to any sparse-data river belonging to the assemblage from which the data-rich rivers are a representative sample.

terms of modelling, exchangeability translates into independent identical distribution (iid) of the θ_i 's. If covariates informative about the θ_i 's variations are available, then exchangeability can still be assumed, conditionally on the covariate. In practice, it is not enough to know that a given variable influences the SR relationship (from some experimental or detailed single site studies). To take advantage of this knowledge, we must be able to measure the covariates on every river of interest, i.e. not only for the data-rich rivers, and also model the nature of the link between the covariates and the θ_i 's. It is clear that these two conditions limit the number of covariates which can be used in practice, especially if we are interested in making inferences for many sparse-data rivers which we know very little about.

Treating the rivers as exchangeable in their SR parameters implies we consider the set of monitored rivers available as a representative sample from a broad assemblage about which we want to make inferences. It means this assemblage must be defined and ideally monitored rivers should have been picked at random. If we know they have been, intentionally or not, selected on criteria which can be influential on the θ_i 's, we should incorporate these criteria as covariates in the analysis, and ensure that the sample of monitored rivers widely covers the range of possibilities regarding these criteria. If these conditions are not met, we are at risk of having a biased view of the broad assemblage of interest (Hilborn and Liermann, 1998).

Data

We define the assemblage of exchangeable rivers of interest to our analysis as the salmon rivers located in western Europe and under the influence of the Gulf Stream. This covers a broad area including Spain, France, UK, Ireland, Norway, the western coast of Sweden and the south-western coast of Iceland. It is known that depending on whether they benefit or not from the warming influence of the Gulf Stream, the rivers differ markedly in their salmon

population dynamics (Scarnecchia, 1984; Scarnecchia *et al.*, 1989; Antonsson *et al.*, 1996).

For semelparous species, it is most useful to express both S and R variables in the same unit (Hilborn and Walters, 1992) because it allows to readily derive estimates of management related parameters from SR analysis, such as MSY, S^* or h^* (i.e., the stock producing MSY, or the exploitation rate at MSY). Although there are exceptions, Atlantic salmon can be treated as a semelparous species, i.e. negligible numbers of adults spawn more than once (Mills, 1989). Because the eggs represent the end product of a generation and the starting point of the next, both stock and recruitment are expressed in terms of eggs. Recruitment was assessed back to the coast, at the time when adults aim at their natal rivers, and prior to any homewater fishery. Removals in distant water fisheries (e.g. at Greenland and Faroes) were considered as a source of random variation in recruitment, as limited information was available to account for the effect of those losses. Measurement errors are not taken into account.

Egg-to-egg SR series can be obtained from monitored rivers, i.e. any river where at least the adult returns and the fisheries are surveyed. Rivers inhabited mainly by sea-trout – holding a comparatively small salmon population – and rivers on which significant stocking programs are conducted without information to apportion their contribution to adult returns were not considered. In addition, only SR series with at least six SR data points since the year of birth 1985 were retained for the analysis. The limitation to the most recent cohorts (after 1985) aims at obtaining Biological Reference Points relevant to the current status of the stocks (see also the more sophisticated time tapering technique proposed by Shepherd (1997)). Non-stationarity in SR relationships is a well-known problem and older data often do not reflect current conditions (Walters and Korman, 2001). Preliminary analysis (not reported here) of the longest SR series available revealed an overall drop in recruitment around the mid-1980s. Data from 13 rivers were retained for analysis (Table 1; Figure 2), ranging from

Table 1. The location, size and SR time series length of the monitored Atlantic salmon rivers retained for analysis.

River	Country	Latitude (°North)	Riverine wetted area accessible to salmon (m ²)	Number of SR observations
Nivelle	France	43	320 995	12
Oir	France	48.5	48 000	14
Frome	UK (England)	50.5	876 420	12
Dee	UK (England)	53	6 170 000	9
Burrishoole	Ireland	54	155 000	12
Lune	UK (England)	54.5	4 230 000	7
Bush	UK (N. Ireland)	55	845 500	13
Mourne	UK (N. Ireland)	55	10 360 560	13
Faughan	UK (N. Ireland)	55	882 380	11
Girnock Burn	UK (Scotland)	57	58 764	12
North Esk	UK (Scotland)	57	2 100 000	6
Laerdalselva	Norway	61	704 000	8
Ellidaar	Iceland	64	199 711	10

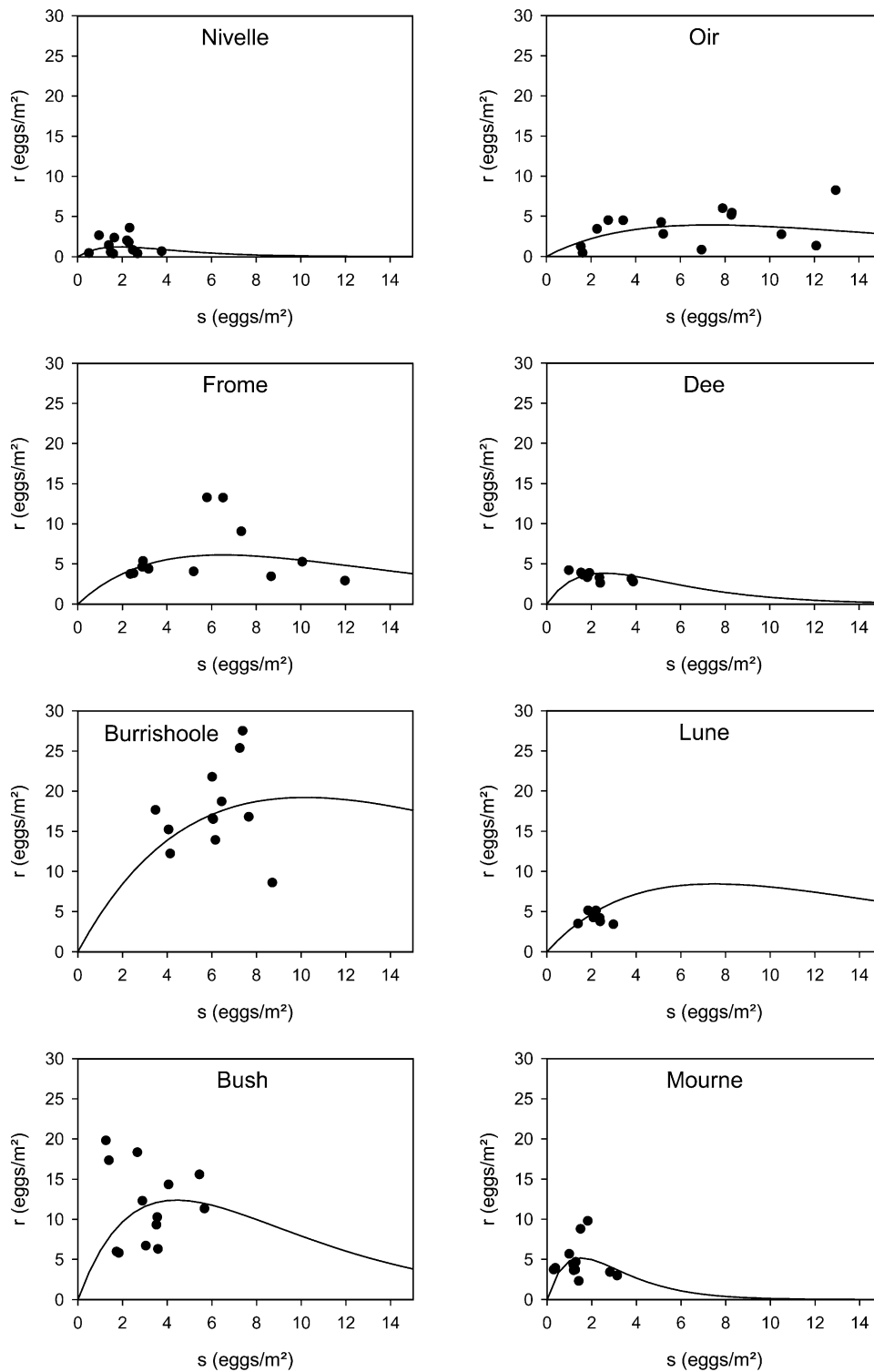


Figure 2. The Atlantic salmon SR series retained for analysis. s and r are the stock and recruitment variables after standardization for river size expressed in eggs per m² of riverine wetted area accessible to salmon. The Ricker curve with parameters set at the median of their posterior distribution are indicated.

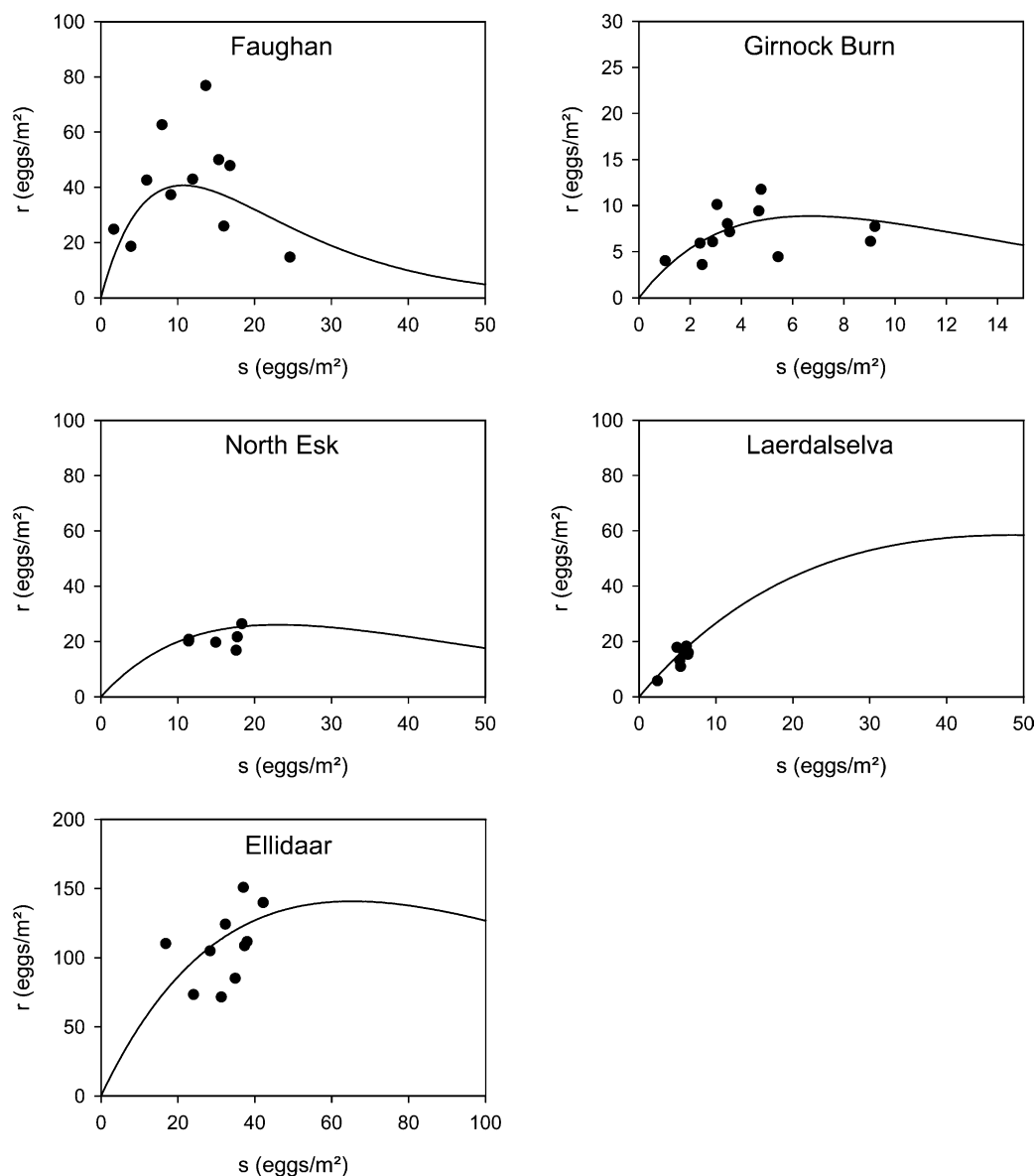


Figure 2 (continued)

the south of France to the south of Iceland. Two are spawning tributaries, the Girnock Burn and the Oir R. The collection and pre-processing procedures used to obtain data ready for analysis are described in detail by Crozier *et al.* (2003).

Among the many covariates we can conceive to explain differences between rivers in their SR parameters, river size is the most evident. The size of a stock is constrained by the size of its river of origin because of the riverine Atlantic salmon ecology. For instance, individuals have a territorial behavior at the juvenile stage and during spawning, and compete for limited spatial resources (Elliott, 2001). Prévost *et al.* (2001) reviewed the many ways of assessing river size

as a limiting factor for salmon production. Currently, the riverine wetted surface area accessible to salmon appears to be the “lowest common denominator” which can be used throughout western Europe. Such a measurement is not yet available for all the salmon rivers, but it could be obtained relatively easily from map based measurements, supported by GIS (Crozier *et al.*, 2003).

Given the very limited information available on the bulk of the salmon rivers, it is difficult to find candidate covariates for explaining variations in SR parameters among rivers. Geographical location is probably the only variable readily accessible. We examined latitude because it shows association with the ecology of Atlantic salmon. For

instance, it is well known that mean smolt age increases with latitude (Metcalf and Thorpe, 1990). Koenings *et al.* (1993) also found a positive latitudinal gradient for smolt-to-adult survival in sockeye salmon (*Oncorhynchus nerka*).

The salmon rivers of England and Wales were used as examples of sparse-data cases to which the transfer of information from the data-rich rivers could be applied. The riverine wetted surface area accessible to salmon and the latitude are known for each of these streams (Anon., 2002).

Model

Within a river, the recruitment process is modeled by means of a Ricker function with independent lognormal process errors. Conditionally on the initial stock level, recruitment observations are assumed to be exchangeable. The following formulation adapted from Schnute and Kronlund (1996) is used:

$$R_{i,j} \sim \text{lognormal}(\log(\text{Ricker}(S_{i,j})), \sigma) \quad (1)$$

$$\begin{aligned} \text{Ricker}(S_{i,j}) = S_{i,j} & (\exp(h_i^*) / (1 - h_i^*)) \\ & \times \exp(-(h_i^* / ((1 - h_i^*) R_i^*)) S_{i,j}) \end{aligned} \quad (2)$$

where

$R_{i,j}$ is the recruitment of the cohort born in year j from the river i ,

$S_{i,j}$ is spawning stock of year $j - 1$ from the river i ,

$\text{Ricker}(S_{i,j})$ is the value of a Ricker function with parameters (h_i^*, R_i^*) at $S_{i,j}$,

σ is the standard deviation of the normal distribution of $\log(R_{i,j})$,

h_i^* is the exploitation rate at MSY for the river i , and

R_i^* is the value of the Ricker function at MSY for the river i .

We assume σ , the dispersion parameter of the lognormal distribution $p(R_{i,j} | S_{i,j}, h_i^*, R_i^*)$, is fixed across rivers. This simplifying hypothesis is motivated by two reasons:

- σ is difficult to estimate on a river by river basis when only short SR series are available. By assuming that σ does not vary among rivers we can cumulate the information from all the SR series.
- σ is not a parameter of interest but a nuisance parameter. Marginal posterior probability distributions of quantities of interest will be obtained by integrating over σ and the existence of several or a single σ will never appear in the end product of the analyses.

Testing of alternative assumptions revealed that the posterior inferences on quantities of interest were not sensitive to the modelling hypothesis about σ (see Appendix A).

Any other SR related parameter or Biological Reference Point can be calculated from h_i^* and R_i^* . The North Atlantic Salmon Conservation Organization (NASCO) recommends the use of the stock level that maximizes the long term average surplus (i.e. MSY) as the standard

Conservation Limit (CL; Potter, 2001). Denoting S_i^* this Biological Reference Point for the river i :

$$S_i^* = (1 - h_i^*) R_i^* \quad (3)$$

The parameters of the Ricker function are assumed to be different among rivers, but drawn from a common probability distribution conditionally on the covariates. The river size must be most influential on R_i^* , i.e. the bigger the river the higher should be R_i^* . This can be translated into

$$R_i^* = r_i^* W_i \quad (4)$$

where

W_i is the riverine wetted area accessible to salmon (m^2), and

r_i^* is the egg recruitment rate per m^2 of riverine wetted area accessible to salmon at MSY.

A preliminary analysis of the monitored rivers SR series was performed on a river by river basis to obtain point estimates of the r_i^* 's and the h_i^* 's (results not reported here). It revealed an increasing latitudinal gradient in the r_i^* 's. Although less evident, the same pattern seemed to exist for the h_i^* 's. The gradient on r_i^* is incorporated in the model by writing:

$$r_i^* \sim \text{lognormal}(\rho_i, \sigma^r) \quad (5)$$

$$\rho_i = A + B \text{lat}_i \quad (6)$$

This log-linear relationship is consistent with the lognormal structure of the SR model used at the within-river level and with the constraint that r_i^* must be positive.

Because h_i^* varies between 0 and 1, it is assigned a beta distribution with parameters α_i and β_i :

$$h_i^* \sim \text{beta}(\alpha_i, \beta_i) \quad (7)$$

The mean of this beta distribution is:

$$\eta_i = \alpha_i / (\alpha_i + \beta_i) \quad (8)$$

We use the following formulation to introduce the dependency of η_i on latitude:

$$\text{logit}(\eta_i) = C + D \text{lat}_i \quad (9)$$

The logit function is used to transform from parameter space $]0, 1[$ to $]-\infty, +\infty[$.

The variance of the h_i^* distribution is:

$$\sigma_i^h = \eta_i(1 - \eta_i) / (\alpha_i + \beta_i + 1) \quad (10)$$

Thus, conditionally on η_i , the variance of the h_i^* distribution is proportional to $1 / (\alpha_i + \beta_i + 1)$. This quantity is a scale parameter directly connected to the "sample size" of the beta distribution, the latter having an intuitive meaning (Congdon, 2001). It is also independent from the mean of the beta distribution, i.e. η_i carries no information about it. We assume this scale parameter is constant across rivers and denote it as

γ . This assumption is equivalent to that made for the distribution of r^*_i (Equation (5)) where σ^r is fixed across rivers.

A, B, C, D, σ^r and γ are called hyper-parameters, i.e. the parameters which, conditionally on the covariates, determine the distribution of the parameters h^*_i and R^*_i . Denoting I the number of monitored rivers considered, the model has $2I + 7$ parameters: two SR parameters for each monitored river plus seven free parameters, the hyper-parameters and the residual dispersion of recruitment σ , which do not depend on any other parameter.

To complete the setting of a full probability model, prior probability distributions are assigned to the free parameters (Table 2). Little informative and independent priors are chosen inspired from recommendations of Box and Tiao (1973). For σ and σ^r the standard non-informative but improper prior for standard deviation parameters, i.e. $p(\sigma) \propto 1/\sigma$ and $p(\sigma^r) \propto 1/\sigma^r$, is approximated by a proper gamma law with shape and scale parameters equal to 0.001 and 1000 (Gelman et al., 1995). Uniform distributions are assigned to the regression coefficients A, B, C and D. Because γ ranges by construction between 0 and 1, it is assigned a beta prior. γ is also a scale parameter. We

choose a beta (0.001,1) prior as a proper distribution which approximates the standard non-informative inverse prior for scale parameters.

Bayesian treatment

The Bayesian treatment of the model amounts to deriving the joint posterior distribution of all the parameters conditionally on the observed data, $p(\theta, \Theta, \sigma | \text{SR}, W, \text{lat})$ where

$$\theta = (\theta_1, \dots, \theta_i, \dots, \theta_I) \quad \text{with} \quad \theta_i = (h^*_i, R^*_i),$$

$$\Theta = (\gamma, A, B, C, D, \sigma^r),$$

SR is the set of SR series from the monitored rivers,

$$W = (W_1, \dots, W_i, \dots, W_I),$$

$$\text{lat} = (\text{lat}_1, \dots, \text{lat}_i, \dots, \text{lat}_I).$$

The joint posterior $p(\theta, \Theta, \sigma | \text{SR}, W, \text{lat})$ and any marginal distribution were approximated using MCMC sampling (Gelman et al., 1995). Gibbs sampling was applied by means of the Winbugs software (Spiegelhalter et al., 2000). Convergence of MCMC sampling toward the target distribution was verified in two steps. We first ran simultaneously four chains with contrasted starting values. We checked their mixing by means of the Gelman–Rubin statistic as modified by Brooks and Gelman (1998), a tool included in Winbugs. Good mixing was reached for all the model parameters after 10 000 iterations. Based on this first step, the first 10 000 iterations of a single chain were discarded and a 10 000 sample was then obtained by carrying on sampling for 100 000 additional iterations and retaining one draw every ten. The stationary of the resulting Markov Chain has been verified for all the free parameters, i.e. A, B, C, D, γ and σ^r , by means of the Geweke (1992) diagnostic as implemented by the R-CODA software (Best et al., 1995). All posterior distributions presented thereafter are approximated by means of this 10 000 sample.

Transfer of information and posterior predictions in sparse-data situations

The distribution of ultimate interest in our case is the posterior predictive of the SR parameters for a new river with no SR data

$$p(\theta_{\text{new}} | \text{SR}, W, \text{lat}, W_{\text{new}}, \text{lat}_{\text{new}}) \quad (11)$$

where

$\theta_{\text{new}} = (h^*_{\text{new}}, R^*_{\text{new}})$, the SR parameters for a new river with no SR data,

W_{new} is the riverine wetted area accessible to salmon of the new river with no SR data, and

lat_{new} is the latitude of the new river with no SR data.

Table 2. Summary of the main parameters and quantities of interest of the model. The prior distributions are indicated only for the free parameters.

Parameter	Definition	Prior distribution
σ	Equation (1)	Gamma (0.001,1000)
Hyper-parameters		
A	Equation (6)	Uniform
B	Equation (6)	Uniform
σ^r	Equation (5)	Gamma (0.001,1000)
C	Equation (9)	Uniform
D	Equation (9)	Uniform
γ	Equation (10) and following section	Beta (0.001,1)
Parameters dependent on hyper-parameters		
r^*_i	Equations (4)–(6)	
h^*_i	Equations (2) and (7)	
Parameters dependent on other parameters		
R^*_i	Equations (2) and (4)	
S^*_i	Equation (3)	
s^*_i	Equation (14)	
Predicted parameters in the absence of SR data		
R^*_{new}	Equivalent to R^*_i	
r^*_{new}	Equivalent to r^*_i	
h^*_{new}	Equivalent to h^*_i	
S^*_{new}	Equivalent to S^*_i	
s^*_{new}	Equivalent to s^*_i	
CL_{reg}	Sum of S^*_{new} 's and S^*_i 's for a region	

This probability distribution can be written as:

$$p(\theta_{\text{new}}|\text{SR}, W, \text{lat}, W_{\text{new}}, \text{lat}_{\text{new}}) \\ = \int p(\theta_{\text{new}}|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})p(\Theta|\text{SR}, W, \text{lat}) d\Theta \quad (12)$$

The posterior predictive is thus obtained by averaging (i.e. integrating) $p(\theta_{\text{new}}|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$ over the posterior distribution of the hyper-parameters. $p(\theta_{\text{new}}|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$ is known. It is the product of two conditionally independent distributions $p(R_{\text{new}}^*|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$ and $p(h_{\text{new}}^*|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$. According to Equations (4)–(6), $p(R_{\text{new}}^*|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$ is a lognormal distribution which dispersion parameter is σ^2 and which location parameter is a function $W_{\text{new}}, \text{lat}_{\text{new}}, A$ and B . According to Equations (7)–(9), $p(h_{\text{new}}^*|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$ is a beta distribution which scale parameter is γ and which location parameter is a function of $\text{lat}_{\text{new}}, C$ and D .

The posterior of the hyper-parameters, $p(\Theta|\text{SR}, W, \text{lat})$, is the distribution through which the SR information coming from the data-rich rivers is transferred to any sparse-data river. It can be expressed as:

$$p(\Theta|\text{SR}, W, \text{lat}) \\ = p(\Theta) \int \dots \int \prod [p(\theta_i|\Theta, W_i, \text{lat}_i)p(\text{SR}_i|\theta_i) d\theta_i] \quad (13)$$

In this expression we see that the information provided by each of the SR_i series is incorporated through the likelihood $p(\text{SR}_i|\theta_i)$ of the parameters of the river i . In this way, the information coming from each data-rich river is judiciously weighed according to how informative it is about the SR related parameters.

Once a posterior sample of Θ has been obtained by MCMC, it is easy to get a sample from $p(\theta_{\text{new}}|\text{SR}, W_{\text{new}}, \text{lat}_{\text{new}})$ for any sparse-data river given its associated W_{new} and lat_{new} values. For any sampled value of Θ , values of R_{new}^* and h_{new}^* can be drawn independently from their known distributions $p(R_{\text{new}}^*|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$ and $p(h_{\text{new}}^*|\Theta, W_{\text{new}}, \text{lat}_{\text{new}})$. Sampling of the posterior predictive distribution of the NASCO CL, noted as S_{new}^* , is then straightforward: an S_{new}^* value is associated to each draw of θ_{new} through Equation (3).

For providing scientific advice for the management of mixed stock fisheries, CLs determined at an aggregated regional level are most useful. Regional CLs are key elements in the procedures used at the International Council for the Exploration of the Sea (ICES) to elaborate scientific advice in response to questions from NASCO (Potter, 2001). A regional CL (CL_{reg}) can be defined as the sum of all the river CLs of a given region. The posterior distribution of CL_{reg} can be denoted as:

$$p(\text{CL}_{\text{reg}}|\text{SR}, W, \text{lat}, W_{\text{reg}}, \text{lat}_{\text{reg}}) \quad (14)$$

where

SR is the set of SR series from the monitored rivers,
 lat_{reg} is the vector of latitudinal positions of the rivers of the region of interest, and

W_{reg} is the vector of riverine wetted areas of the rivers of the region of interest.

The CLs of the sparse-data rivers are independent conditionally on the hyper-parameters Θ . In other words, they depend on the SR data collected from the data-rich rivers only through the hyper-parameters. Therefore, it is straightforward to get a sample of CL_{reg} values to approximate the CL_{reg} posterior predictive distribution. It amounts, for each draw of Θ in its posterior distribution, to successively and independently draw an S_{new}^* value for each river in the region of interest, and then calculate the sum of these river CLs to generate a CL_{reg} value.

Because CL_{reg} is a sum of variables with (conditionally) independent distributions, the precision of the posterior distribution of CL_{reg} will be improved compared to that of an individual river S_{new}^* . Even imprecise SR related parameters estimates obtained for sparse-data rivers could then be valuable information when aggregated at a regional level. This approach is applied to the salmon rivers of England and Wales.

Results

Monitored rivers vary widely in size (Table 1), from 48 000 (Oir R.) to 10.4 million m^2 (Mourne R.) of riverine wetted area accessible to salmon. Comparisons between SR data series can be made after removing the river size effect by looking at SR rates per m^2 of riverine wetted area accessible to salmon, denoted as s and r . Contrast in s is poor for various rivers as shown by the width of the inter-quartile range (Figure 3). Within-river contrast in r is similar to that of s . There is a tendency of increasing r when moving northward. Such a pattern is much less evident for s .

The posterior probability that $B < 0$ is null while that of $D < 0$ is 0.05. Our choice of introducing latitude as a covariate for explaining variations between rivers in both r^* and h^* is thus validated. The increasing latitudinal gradient in r^* is obvious in Figure 4. h^* is often poorly estimated (Figure 5) and the latitudinal gradient is less evident.

As for R^* and r^* (Equation (4)), we define:

$$s^* = S^*/W \quad (15)$$

the egg deposition rate per m^2 of riverine wetted area accessible to salmon at MSY, which is more easily comparable between rivers as it is standardized for the river size variations. The posterior distributions of s^* for the monitored rivers reveal (Figure 6):

- considerable within-river uncertainty in some cases despite SR data being available (e.g. the Lune R. and the Laerdalselva R.),
- significant variations among rivers, even within a relatively narrow latitudinal range (e.g. the Bush R., the Mourne R. and the Faughan R., all located in Northern Ireland), and
- an increasing trend with latitude.

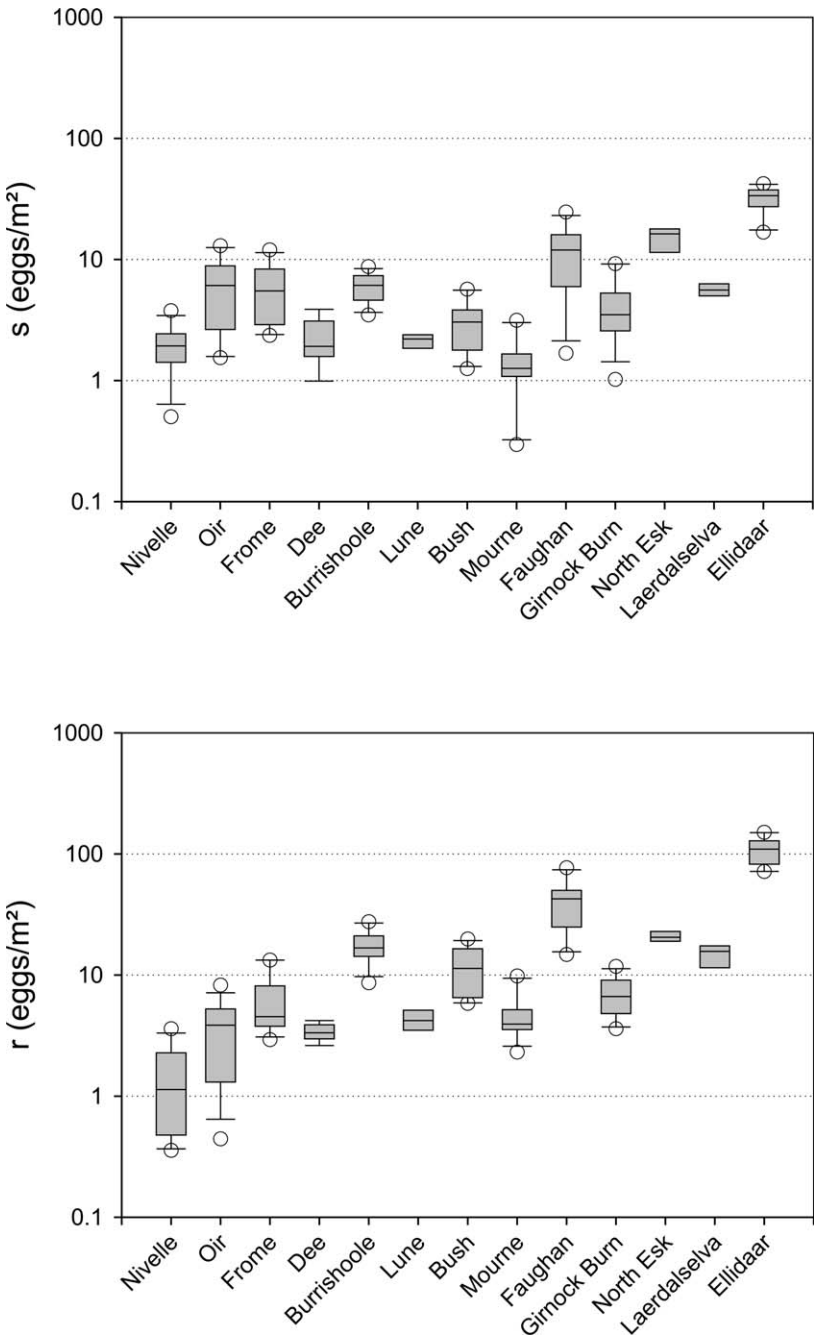


Figure 3. Box plots of the stock (upper panel) and recruitment (lower panel) series from the data-rich rivers. Raw data are standardized for river size and expressed in eggs per m² of riverine wetted area accessible to salmon. Each box plot displays on a log scale the 10th, 25th, 50th, 75th, 90th percentiles and all values outside the 10th to 90th percentiles interval. Rivers are ordered by increasing latitude.

The distributions of ultimate interest are the posterior predictive distributions, which represent our uncertainty/knowledge for sparse-data rivers without SR observations. The marginal posterior predictive distribution of r^*_{new} , h^*_{new} and s^*_{new} at various latitudes covering the salmon

distribution range in the NEAC area (45°, 50°, 55°, 60° and 65° North) indicate that when moving northward, salmon stocks can sustain higher exploitation rates, can produce higher recruitment at MSY, but at the same time should be set at higher CLs (Table 3, Figure 6). All the

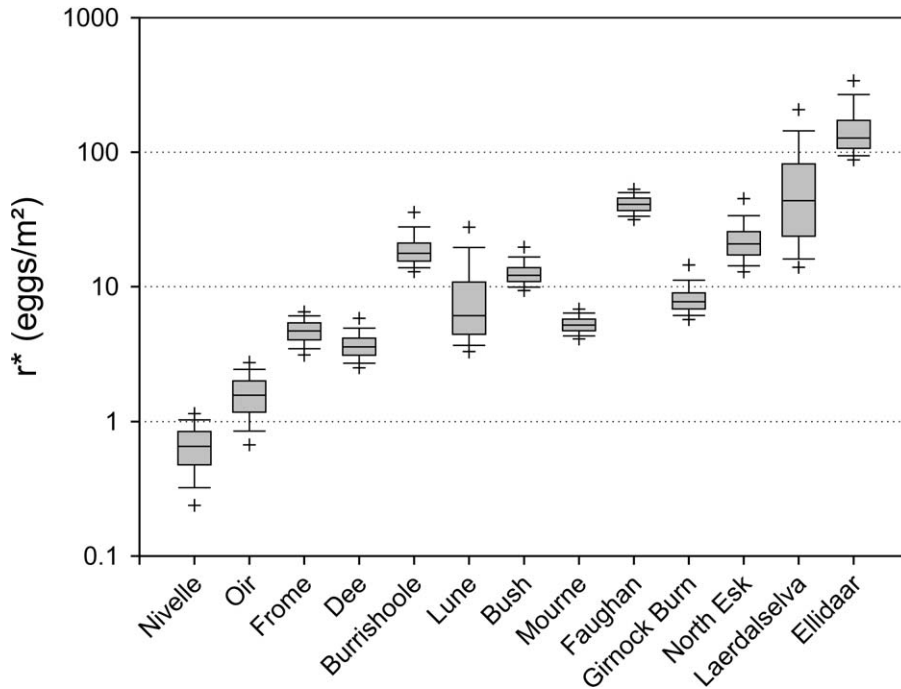


Figure 4. Box plots of the posterior distributions of r^* (the egg recruitment rate per m^2 of riverine wetted area accessible to salmon at MSY) for the data-rich rivers. Each box plot displays on a log scale the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles. Rivers are ordered by increasing latitude.

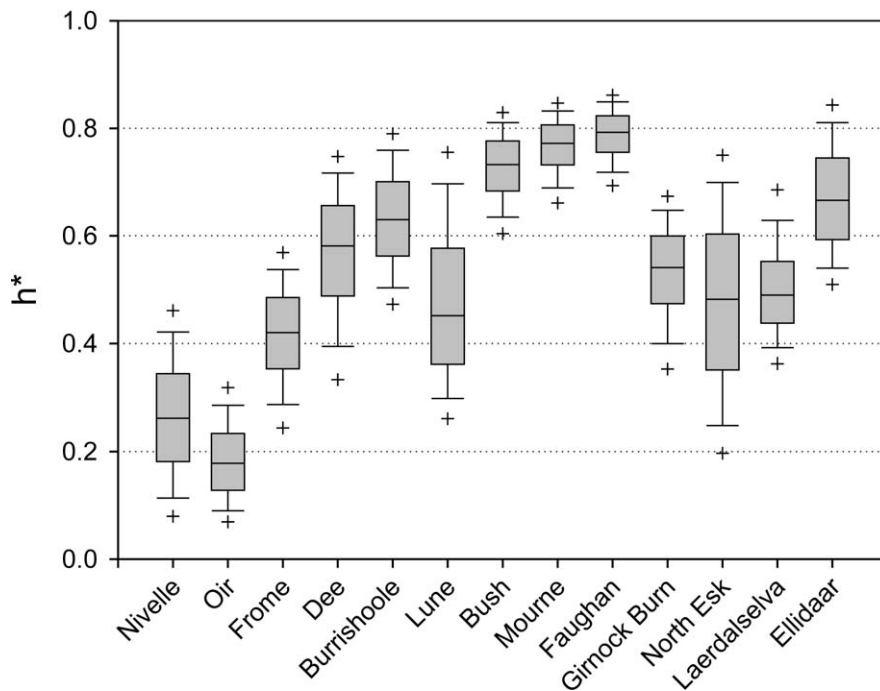


Figure 5. Box plots of the posterior distributions of h^* (the exploitation rate at MSY) for the data-rich rivers. Each box plot displays the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles. Rivers are ordered by increasing latitude.

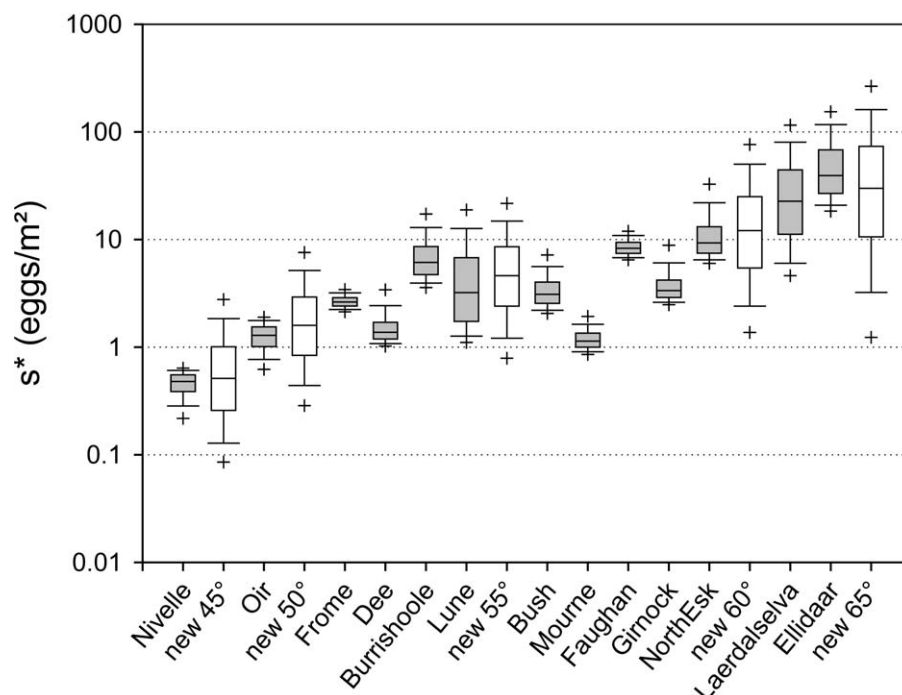


Figure 6. Box plots of the distributions of s^* (egg deposition rate per m^2 of riverine wetted area accessible to salmon at MSY). Gray: posteriors for the data-rich rivers. White: predictive posteriors for any sparse-data river located at 45°, 50°, 55°, 60° and 65° North latitude. Each box plot displays on a log scale the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles. Rivers are ordered by increasing latitude.

posterior predictive distributions are very wide, suggesting that there is great uncertainty in the spawning stock, the recruitment and the exploitation rates at MSY for a sparse-data river.

There are 70 salmon rivers in England and Wales located between 55.5° and 50.5° North. They vary in size from 80 000 to 14 million m^2 of riverine wetted area accessible to salmon (median 610 000; Anon., 2002), thus covering a size range comparable with that of the data-rich rivers. We compare the posterior predictive distribution of CL_{reg} for England and Wales with that of its river components. In order to make this comparison independently from a size effect, the average egg deposition rate corresponding to the England and Wales CL_{reg} is calculated by dividing CL_{reg} by the sum of the riverine wetted area accessible to salmon of all the England and Wales salmon rivers. The posterior predictive distribution of this average rate is more precise than that of the s^*_{new} of its individual river components (Figure 7). The average egg deposition rate corresponding to the England and Wales CL_{reg} is less precisely estimated than s^* for the Frome and Dee Rivers, two England and Wales data-rich rivers. This does not hold for the Lune River, because the SR data are essentially uninformative about s^* due to a lack of contrast in the stock variable (Figure 3). For the Lune R., the posterior of s^* is close to the predictive posterior of s^*_{new} , i.e. in the absence of SR data (Figure 6). Despite the variance reduction due to the

averaging effect when dealing with broad regions such as England and Wales, the posterior predictive distribution of CL_{reg} for England and Wales remains relatively wide. The 80% probability interval (10th to 90th percentile interval) is 286–776 million eggs.

Discussion

A chief merit of the hierarchical approach we implemented in the Bayesian hierarchical SR analysis, is that it provides a common and consistent framework for the joint analysis of several monitored stocks data sets. This is a step towards making extensive use of detailed stock monitoring programs for improving management advice. The hierarchical modelling allows the transfer of information from the data-rich stocks towards the others while acknowledging the existence of differences between stocks. These differences are a source of uncertainty when extrapolating to sparse-data situations. For Atlantic salmon, this is an improvement over current practices and methods proposed for setting and transporting Biological Reference Points. These completely ignore uncertainty (Elson, 1957, 1975; Potter, 2001), or when accounting for some (Chaput *et al.*, 1998; Caron *et al.*, 1999) unrealistically assume that there is no variation between rivers, or when acknowledging some differences between rivers still rely on a single monitored river SR data set (Prévost and Porcher, 1996; Milner *et al.*, 2000).

Table 3. Summary statistics (mean, standard deviation, 10th percentile, median, 90th percentile) of the marginal posterior distributions for the free parameters of the hierarchical SR model and of the posterior predictive distribution of h^*_{new} and r^*_{new} (the egg recruitment rate per m^2 of riverine wetted area accessible to salmon at MSY) at various latitudes.

		Mean	Std. dev	10th	Median	90th
Posterior distributions of free parameters						
	σ	0.48	0.03	0.44	0.48	0.52
Hyper-parameters						
	A	-11.94	3.11	-15.81	-11.85	-8.21
	B	0.26	0.06	0.19	0.26	0.33
	σ^r	0.79	0.22	0.56	0.75	1.08
	C	-4.74	3.10	-8.65	-4.63	-0.91
	D	0.09	0.06	0.02	0.09	0.16
	γ	0.14	0.06	0.07	0.13	0.22
Posterior predictive distributions						
Latitude ($^{\circ}$ North)	Parameter					
45	h^*_{new}	0.34	0.21	0.08	0.32	0.63
50	h^*_{new}	0.44	0.20	0.18	0.44	0.71
55	h^*_{new}	0.55	0.19	0.29	0.56	0.80
60	h^*_{new}	0.65	0.20	0.38	0.68	0.90
65	h^*_{new}	0.73	0.20	0.45	0.77	0.96
45	r^*_{new}	1.41	3.29	0.23	0.83	2.77
50	r^*_{new}	4.70	7.74	1.00	3.05	9.14
55	r^*_{new}	16.84	30.44	3.91	11.14	30.89
60	r^*_{new}	68.54	279.42	13.25	40.17	127.90
65	r^*_{new}	279.10	785.28	40.52	149.50	535.00

The hierarchical framework is also flexible; it potentially can accommodate any type of model representing single stock dynamics. The Bayesian hierarchical SR analysis as presented here does not address two of the well-known pitfalls of SR analysis (Walters and Korman, 2001), measurement errors on the SR variables and the time series structure of the SR data sets (Meyer and Millar, 2000; Rivot *et al.*, 2001; Schnute and Kronlund, 2002), but there is no impediment to incorporating more realistic or complex models in the near future. By allowing the derivation of posterior predictive distributions, the hierarchical Bayesian setting provides a way of describing probabilistically our knowledge about “the states of nature” even in sparse-data situations, thus opening the door for decision analysis along the lines presented by Punt and Hilborn (1997) and Hilborn (2001).

However, posterior predictive inferences may be very imprecise. This point is well exemplified in our case study. Recruitment is known to be a highly variable process (Fogarty *et al.*, 1991) and thus SR related parameters could not be estimated precisely with short SR series (Figures 4–6). This is especially true when the contrast in stock size is poor (Figure 3). In addition, many features, other than riverine wetted area accessible to salmon and latitude, can cause variations in the recruitment process between rivers. We should not expect to be able to provide precise estimates of SR related parameters on the basis of the limited data used in our analysis and more generally from the scarce information currently available for the vast majority of the salmon rivers.

Given the narrower posterior distributions obtained for Biological Reference Points on the data-rich rivers, it might be tempting to directly extrapolate those distributions across neighboring rivers rather than using posterior predictive distributions incorporating information from far away stocks. The joint treatment of the several SR series presented herein is a strong warning against such practices. It clearly shows that, even in a narrow geographical range, s^* can vary widely (see for instance the three monitored rivers from Northern Ireland; Figure 6). Over-reliance on local data-rich stocks should be dismissed as it can lead to a major underestimation of our uncertainty about management parameters for sparse-data rivers.

The aggregation of several stocks under a regional complex improves the precision of the posterior predictive inferences. When several stocks are managed jointly, even imprecise knowledge about each component of the aggregate can be valuable. This is well exemplified by our case study on the salmon rivers of England and Wales. The location of the posterior predictive distribution of the egg deposition rate corresponding to CL_{reg} for England and Wales is very close to that of s^*_{new} at 55° North (Figure 7), despite the fact the salmon rivers of England and Wales are well spread between 50° and 55° North. Summing over many rivers CLs drawn from distributions which do not exclude very high values tend to exclude low values for the aggregated CL_{reg} . Consequently, deriving a point estimate regional CL by summing the most likely river CL values is not equivalent to the most likely value of a sum of CLs. Ignoring the wide and skewed uncertainty in s^*_{new} would

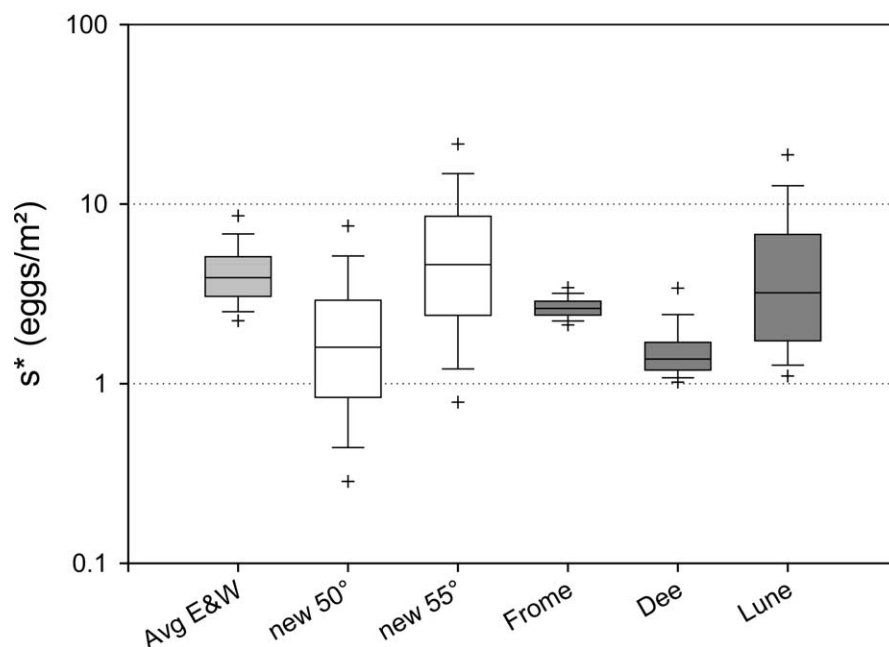


Figure 7. Box plots of distributions of egg deposition rate per m^2 of riverine wetted area accessible to salmon: (i) the posterior predictive of the average egg deposition rate corresponding to CL_{reg} for England and Wales (in gray); (ii) the posterior predictive of s^*_{new} for any river located at 50° and 55° North (in white) and (iii) the posterior of s^* for the data-rich rivers located in England and Wales (in dark gray). Each box plot displays on a log scale the 5th, 10th, 25th, 50th, 75th, 90th and 95th percentiles.

lead to an underestimation of the regional CL for England and Wales. In the context of providing management advice, acknowledging uncertainty associated to each stock component matters when setting a CL for a regional aggregate.

The introduction of the river size and latitude covariates for explaining between stock variations resulted in a significant gain of precision for the posterior predictive inferences. Without these two covariates, the variations they generate between stocks would have been treated as random noise, thus blurring all posterior predictions for sparse-data stocks. Nevertheless, very limited information is incorporated in our Bayesian hierarchical SR analysis. Additional knowledge on the monitored rivers – e.g. sympatric brown trout (*Salmo trutta*), presence of lakes, and habitat quality – is ignored. There is certainly ample room for improvement by taking greater advantage of the scientific knowledge available about factors influencing SR processes at the river level. But it is not enough to identify a factor that influences the stock dynamics. The operational translation of this knowledge in the implementation of a broad scale analysis also requires that the relevant data are made available over a broad geographical range and in a standardized manner. Moreover, the nature of the link between a factor and the population dynamics processes must be translated into a formal relationship between a relevant covariate and the SR parameters. Taking the example of the presence of trout and the recruitment process in salmon, we know there are interactions between juvenile salmon and trout in streams (Heggenes *et al.*,

1999). However, not only are data on trout populations not available for all the European salmon rivers, but the identification of a relevant covariate and the modalities of its effect on SR parameters are also unclear. The presence of trout is thus only one of the many factors which causes random (unpredictable) variations in the SR relationship among rivers. Such difficulties should not prevent trying to identify relevant and practically usable covariates to improve the posterior predictive inferences. In our case, the introduction in the Bayesian hierarchical SR analysis of an indication of slope along the stretches of a river colonized by salmon could be a valuable track to explore. Slope is related to the quality of habitat for juveniles (Amiro, 1993) and it could be measured over a broad geographical range by means of GIS. Data collection regarding candidate covariates should at first be limited to the monitored stocks in order to assess their potential in explaining differences between stocks in their dynamics.

Covariates could also be used to assess the risk of bias in the inferences made from the available set of data-rich stocks. This set is regarded as a representative sample of a larger assemblage (e.g. the salmon rivers of Europe influenced by the Gulf stream in our case study) covering the full range of possible values of the covariates. In most cases this remains arguable. The set of data-rich salmon rivers we analyzed contain only one lacustrine system (the Burrish-oolle) and Scandinavia is under-represented (Table 1). The little or non-represented categories of rivers might exhibit consistent differences in their SR relationship compared to

the data-rich rivers. In the absence of data we have limited capability to assess the associated potential bias. Taking the presence of lakes as an example, we are left with the following choice. If we assume the presence of lakes has a consistent positive (or negative) effect on salmon recruitment, not accounted for by the *W* variable as used in Equation (4), then we should discard the Burrishoole R. and restrict the analysis and the subsequent inferences to riverine-only systems. Alternatively, if we assume lakes have an unpredictable random effect on the recruitment, then we can use the current set of monitored rivers to make inferences about all the rivers. This latter option is not a contradiction to the results of the Bayesian hierarchical SR analysis reported herein: the Burrishoole R. does not appear as an outlier (Figures 4–6) when compared to the riverine-only systems.

Before using the Bayesian hierarchical SR analysis or any other Bayesian hierarchical modelling approach for management advice, a decision must be explicitly made, communicated and discussed about the exact contour of the assemblage of stocks from which the data-rich stocks are seen as a representative sample. Gurevitch *et al.* (2001) provide statistical guidelines for the combination of ecological data sets in the perspective of a meta-analysis. The definition of the assemblage of stocks which we model as exchangeable units, conditionally on the covariates, remains the most influential choice to be made when attempting to transfer information from data-rich to sparse-data situations. Longer SR time series would help in reducing the uncertainty of population dynamics related parameters, as far as the contrast in the independent/control variables (i.e. spawning stock in our case study) increases with time. Walters and Korman (2001) even suggest to deliberately manipulate some stocks through management experiments to get more informative data. However, whatever the precision of our posterior inferences in data-rich situations, it is the large between-stocks residual variation left after accounting for the effect of the readily available covariates (i.e. river size and latitude in our case study) which impedes precise posterior predictions in sparse-data situations. For improving precision, the introduction of additional covariates informative about the population dynamics related parameters of interest is required.

Acknowledgements

This work was undertaken within the SALMODEL project, a EU funded Concerted Action (Contract No: QLK5-CT1999-01546). We thank the two anonymous referees whose comments helped to improve the content of the paper.

References

- Amiro, P. G. 1993. Habitat measurement and population estimation of juvenile Atlantic salmon (*Salmo salar*). In *Production of Juvenile Atlantic Salmon, Salmo salar*, in Natural Waters, pp. 81–97. Ed. by R. J. Gibson, and R. E. Cutting, Canadian Special Publication of Fisheries and Aquatic Sciences, 118.
- Anon. 2002. Annual Assessment of Salmon Stocks and Fisheries in England and Wales. CEFAS/Environment Agency, Lowestoft/Cardiff, 59 pp.
- Antonsson, T., Gudbergsson, G., and Gudjonsson, S. 1996. Environmental continuity in fluctuation of fish stocks in the North Atlantic Ocean, with particular reference to Atlantic salmon. *North American Journal of Fisheries Management*, 16: 540–547.
- Berger, J. O. 1985. *Statistical decision theory and Bayesian analysis*. Springer Series in Statistics. 2nd edition. Springer-Verlag, New York. 617 pp.
- Best, N. G., Cowles, M. K., and Vines, S. K. 1995. CODA Manual Version 0.30. MRC, Biostatistics Unit, Cambridge. 41 pp.
- Brooks, S. P., and Gelman, A. 1998. Alternative methods for monitoring convergence of iterative simulations. *Journal of Computational and Graphical Statistics*, 7: 434–455.
- Box, G. E. P., and Tiao, G. C. 1973. *Bayesian inference in statistical analysis*. Wiley Classics Library, John Wiley and Sons, New York. 588 pp.
- Caron, F., Fontaine, P.-M., and Picard, S.-E. 1999. Seuil de conservation et cible de gestion pour les rivières à saumon (*Salmo salar*) du Québec. Faune et Parcs Québec, Direction de la faune et des habitats, Québec. 45 pp.
- Chaput, G., Allard, J., Caron, F., Dempson, J. B., Mullins, C. C., and O'Connell, M. F. 1998. River-specific target spawning requirements for Atlantic salmon (*Salmo salar*) based on a generalized smolt production model. *Canadian Journal of Fisheries and Aquatic Sciences*, 55: 246–261.
- Chen, D. G., and Holtby, L. B. 2002. A regional meta-model for stock recruitment analysis using an empirical Bayesian approach. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 1503–1514.
- Congdon, P. 2001. *Bayesian Statistical Modelling*. Wiley Series in Probability and Statistics. Wiley, Chichester. 529 pp.
- Crozier, W. W., Potter, E. C. E., Prévost, E., Schön, P.-J., and Ó Maoiléidigh, N. (eds) 2003. A coordinated approach towards the development of a scientific basis for management of wild Atlantic salmon in the North-East Atlantic. Report of Concerted Action QLK5-CT1999–01546 (SALMODEL). 420 pp.
- Dorn, M. W. 2002. Advice on west coast rockfish harvest rates from Bayesian meta-analysis of stock–recruit relationship. *North American Journal of Fisheries Management*, 22: 280–300.
- Elliott, J. M. 2001. The relative role of density in the stock–recruitment relationship of salmonids. In *Stock, Recruitment and Reference Points – Assessment and Management of Atlantic Salmon*, pp. 25–66. Ed. by E. Prévost, and G. Chaput. Hydrobiologie et Aquaculture, INRA, Paris.
- Elson, P. F. 1957. Number of salmon needed to maintain stocks. *Canadian Fish Culturist*, 21: 18–23.
- Elson, P. F. 1975. Atlantic salmon rivers smolt production and optimal escapement. An overview of natural production. In *New England Atlantic salmon restoration conference 1975*, pp. 96–119. Ed. by J. R. Bohne, and L. Sochasky. International Atlantic Salmon Foundation Special Publication Series, 6.
- Fogarty, M. J., Sissenwine, M. P., and Cohen, E. B. 1991. Recruitment variability and the dynamics of exploited marine populations. *TREE*, 6: 241–246.
- Gelman, A., Carlin, J. B., Stern, H. L., and Rubin, D. B. 1995. *Bayesian data analysis*. Chapman and Hall, London. 526 pp.
- Geweke, J. 1992. Evaluating accuracy of sampling-based approaches to the calculation of posterior moments. In *Bayesian Statistics 4*, pp. 169–193. Ed. by J. M. Bernardo, J. Berger, A. P. Dawid, and A. F. M. Smith. Oxford University Press, Oxford.

- Gurevitch, J., Curtis, P. S., and Jones, M. H. 2001. Meta-analysis in ecology. *Advances in Ecological Research*, 32: 199–247.
- Heggenes, J., Bagliniere, J.-L., and Cunjak, R. 1999. Spatial niche variability for young Atlantic salmon (*Salmo salar*) and brown trout (*Salmo trutta*). *Ecology of Freshwater Fish*, 8: 1–21.
- Hilborn, R. 2001. Risk analysis for salmon spawning reference levels. *In* Stock, Recruitment and Reference Points – Assessment and Management of Atlantic Salmon, pp. 177–193. Ed. by E. Prévost, and G. Chaput. Hydrobiologie et Aquaculture, INRA, Paris.
- Hilborn, R., and Liermann, M. 1998. Standing on the shoulders of giants: learning from experience in fisheries. *Reviews in Fish Biology and Fisheries*, 8: 273–283.
- Hilborn, R., and Walters, C. J. 1992. Quantitative fisheries stock assessment. Choice, dynamics and uncertainty. Chapman and Hall, New York/London. 570 pp.
- Koenings, J. P., Geiger, H. J., and Hasbrouck, J. J. 1993. Smolt-to-adult survival patterns of sockeye salmon (*Oncorhynchus nerka*): effects of smolt length and geographic latitude when entering the sea. *Canadian Journal of Fisheries and Aquatic Sciences*, 50: 600–611.
- Liermann, M., and Hilborn, R. 1997. Depensation in fish stocks: a hierarchical Bayesian meta-analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 54: 1976–1984.
- Link, W. A., Cam, E., Nichols, J. D., and Cooch, E. G. 2002. Of BUGS and birds: Markov Chain Monte Carlo for hierarchical modelling in wildlife research. *Journal of Wildlife Management*, 66: 277–291.
- Metcalfe, N. B., and Thorpe, J. E. 1990. Determinants of geographical variation in the age of seaward-migrating salmon (*Salmo salar*). *Journal of Animal Ecology*, 64: 2339–2346.
- Meyer, R., and Millar, R. 2000. State-space models for stock-recruit time series. *In* Bayesian Methods with Applications to Science, Policy, and Official Statistics – Selected Papers from ISBA 2000, pp. 361–370. Ed. by E. I. George. (<http://www.stat.cmu.edu/ISBA/index.html>).
- Mills, D. 1989. Ecology and Management of Atlantic Salmon. Chapman and Hall, London. 351 pp.
- Milner, N. J., Davidson, I. C., Wyatt, R. J., and Aprahamian, M. 2000. The use of spawning targets for salmon fishery management in England and Wales. *In* Management and ecology of river fisheries, pp. 360–372. Ed. by I. G. Cowx. Fishing News Books, Blackwell Scientific Publications, Oxford.
- Myers, R. A. 2001. Stock and recruitment: generalizations about maximum reproductive rate, density dependence and variability using meta-analyses approaches. *ICES Journal of Marine Science*, 58: 937–951.
- Myers, R. A., and 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.
- Myers, R. A., MacKenzie, B. R., Bowen, K. G., and Barrowman, N. J. 2001. What is the carrying capacity for fish in the Ocean? A meta-analysis of population dynamics of North Atlantic cod. *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 1464–1476.
- Myers, R. A., Barrowman, N. J., Hilborn, R., and Kehler, D. G. 2002. Inferring Bayesian priors with limited direct data: application to risk analysis. *North American Journal of Fisheries Management*, 22: 351–364.
- Potter, T. 2001. Past and present use of reference points for Atlantic salmon. *In* Stock, Recruitment and Reference Points – Assessment and Management of Atlantic Salmon, pp. 195–223. Ed. by E. Prévost, and G. Chaput. Hydrobiologie et aquaculture, INRA, Paris.
- Prévost, E., and Chaput, G. (eds) 2001. Stock, Recruitment and Reference Points – Assessment and Management of Atlantic Salmon. Hydrobiologie et Aquaculture, INRA, Paris. 223 pp.
- Prévost, E., and Porcher, J.-P., 1996. Méthodologie d'élaboration de totaux autorisés de captures (TAC) pour le Saumon atlantique (*Salmo salar* L.) dans le Massif Armoricain. Propositions et recommandations scientifiques. GRISAM, Évaluation et gestion des stocks de poissons migrateurs, Document scientifique et technique 1. 18 pp.
- Prévost, E., Chaput, G., and Chadwick, E. M. P. 2001. Transport of stock-recruitment reference points for Atlantic salmon. *In* Stock, Recruitment and Reference Points – Assessment and Management of Atlantic Salmon, pp. 95–135. Ed. by E. Prévost, and G. Chaput. Hydrobiologie et Aquaculture, INRA, Paris.
- Punt, A., and Hilborn, R. 1997. Fisheries stocks assessment and decision analysis: the Bayesian approach. *Reviews in Fish Biology and Fisheries*, 7: 35–63.
- Rivot, E., and Prévost, E. 2002. Hierarchical Bayesian analysis of capture-mark-recapture data. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 1768–1784.
- Rivot, E., Prévost, E., and Parent, E. 2001. How robust are Bayesian inferences based on a Ricker model with regards to measurement errors and prior assumptions about parameters? *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 2284–2297.
- Scarnecchia, D. L. 1984. Climatic and oceanic variations affecting yield of Icelandic stocks of Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, 41: 917–935.
- Scarnecchia, D. L., Isaksson, A., and White, S. E. 1989. Oceanic and riverine influences on variations in yield among Icelandic stocks of Atlantic salmon. *Transactions of the American Fisheries Society*, 118: 482–494.
- Schnute, J. T., and Kronlund, A. R. 1996. A management oriented approach to stock recruitment analysis. *Canadian Journal of Fisheries and Aquatic Sciences*, 53: 1281–1293.
- Schnute, J. T., and Kronlund, A. R. 2002. Estimating salmon stock-recruitment relationships from catch and escapement data. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 433–449.
- Shepherd, J. G. 1997. Prediction of year-class strength by calibration regression analysis of multiple recruit index series. *ICES Journal of Marine Science*, 54: 741–752.
- Spiegelhalter, D. J., Thomas, A., and Best, N. G. 1996. Computation on graphical Bayesian models. *In* Bayesian Statistics 5, pp. 407–425. Ed. by J. M. Bernardo, J. O. Berger, A. P. Dawid, and A. F. M. Smith. Oxford University Press.
- Spiegelhalter, D., Thomas, A., and Best, N. 2000. WinBUGS version 1.3. User Manual. MRC and Imperial College of Science, Technology and Medicine, 34 pp. (<http://www.mrc-bsu.cam.ac.uk/bugs>).
- Su, Z., Adkinson, M. D., and Van Alen, B. W. 2001. A hierarchical Bayesian model for estimating historical salmon escapement and escapement timing. *Canadian Journal of Fisheries and Aquatic Sciences*, 58: 1648–1662.
- Walters, C., and Korman, J. 2001. Analysis of stock-recruitment data for deriving escapement reference points. *In* Stock, Recruitment and Reference Points – Assessment and Management of Atlantic Salmon, pp. 67–94. Ed. by E. Prévost, and G. Chaput. Hydrobiologie et Aquaculture, INRA, Paris.
- Wyatt, R. J. 2002. Estimating riverine fish population size from single- and multiple-pass removal sampling using a hierarchical model. *Canadian Journal of Fisheries and Aquatic Sciences*, 59: 695–706.

Appendix A

In order to assess the sensitivity of our results to the modelling hypothesis about the dispersion parameter of the lognormal distribution $p(R_{ij}|S_{ij}, h_i^*, R_i^*)$, we replaced the assumption of a fixed σ across rivers by the radically contrasted option of different and fully unrelated (independent) σ_i 's. Apart from this change, we reprocessed the data using exactly the same approach and compared the posterior inferences for all the free parameters which govern the quantities of interest in our hierarchical SR model (Tables 3 and A.1). The σ_i 's vary among rivers but their posterior distributions substantially overlap. The 80%

posterior probability intervals derived under the full independence hypothesis are wider than that obtained under the complete pulling hypothesis. The single river 80% probability intervals do not necessarily overlap with that obtained when ignoring variation between rivers. Despite these differences with regard to the nuisance parameter σ , the posterior distributions of the hyper-parameters are little changed. Intermediate assumption about the σ_i 's, such as a hierarchical structure or a latitudinal gradient, were also tested. As expected they produce similar results (not reported here) thus confirming, in our case study, the overall insensitivity of the posterior inferences to the modelling hypothesis about the σ_i 's.

Table A.1. Summary statistics (mean, standard deviation, 10th percentile, median, 90th percentile) of the marginal posterior distributions of the free parameters of the hierarchical SR model under the hypothesis of fully independent σ_i 's dispersion parameters of the lognormal distributions $p(R_{ij}|S_{ij}, h_i^*, R_i^*)$.

River	Parameter	Mean	Std. dev	10th	Median	90th
Nivelle	σ	0.85	0.20	0.63	0.81	1.11
Oir	σ	0.85	0.19	0.64	0.82	1.10
Frome	σ	0.48	0.12	0.36	0.46	0.64
Dee	σ	0.21	0.07	0.14	0.20	0.29
Burrishoole	σ	0.35	0.09	0.25	0.33	0.45
Lune	σ	0.20	0.09	0.12	0.18	0.31
Bush	σ	0.58	0.14	0.43	0.56	0.75
Mourne	σ	0.48	0.13	0.34	0.45	0.64
Faughan	σ	0.48	0.15	0.33	0.44	0.66
Girnock Burn	σ	0.36	0.09	0.26	0.34	0.47
North Esk	σ	0.20	0.08	0.12	0.18	0.30
Laerdalselva	σ	0.21	0.07	0.14	0.19	0.29
Ellidaar	σ	0.29	0.08	0.21	0.27	0.39
	Hyper-parameters					
	A	-11.00	3.10	-14.96	-10.91	-7.23
	B	0.24	0.06	0.17	0.24	0.32
	σ^r	0.79	0.21	0.56	0.75	1.05
	C	-3.84	3.05	-7.73	-3.65	-0.16
	D	0.08	0.06	0.01	0.07	0.15
	γ	0.11	0.06	0.05	0.10	0.19