Borrowing strength from similar stock-recruitment analyses
HBM for Ecological Data, chapter 9.2

Éric Parent & Étienne Rivot
AgroParisTech, UMR518, Dpt of Maths, Paris, France
AgroCampus Ouest, UMR0985, Dpt of Ecology, Rennes, France
Eric.Parent@agroparistech.fr
Etienne.Rivot@agrocampus-ouest.fr

Monday, the 30th of June 2014

Borrowing strength from data-poor to data-rich units

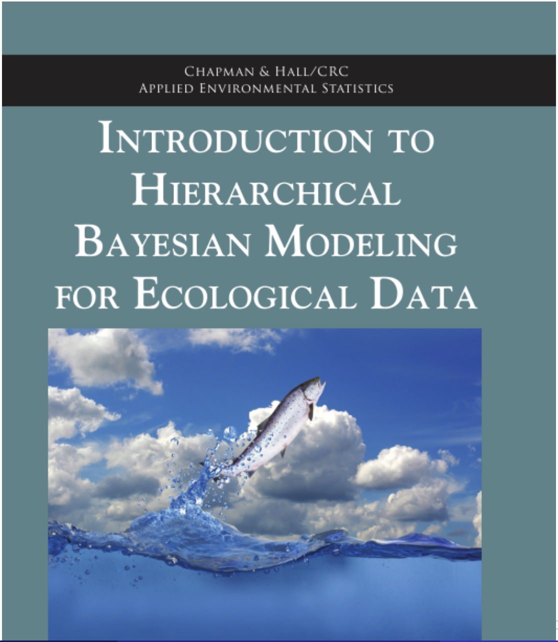
Hierarchical (*multilevel* or *random effect*) models assume that the dataset being analyzed consists of a *hierarchy* of different *groups* within which records look more alike than between groups.

- SR hierarchical analysis for 13 rivers in Europe : from data-rich to sparse situations.
- Latitude and riverine wetted area accessible to salmon as covariates.
- Hierarchical structure to organize the transfer of information between different units.

In renewable resources management models :

- *Cohort effects* in correlated or familial survival data as opposed to individual behaviors within a group ;
- *Site effects* in meta-analyzes or in *spatially* structured phenomena.

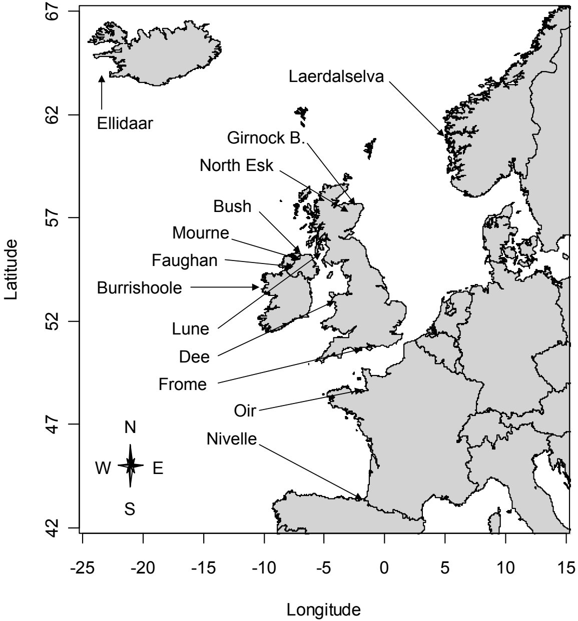
This presentation is extracted from chapter 9



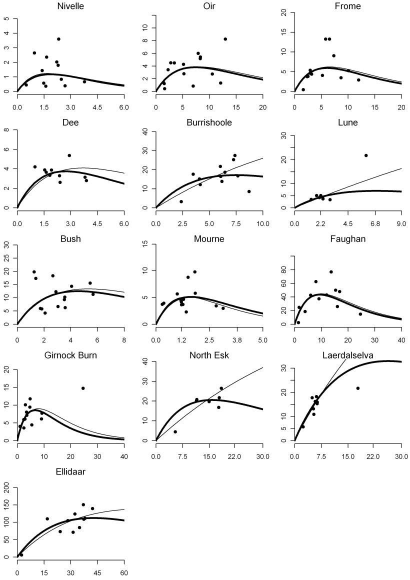
Data for Hierarchical stock-recruitment analysis

index rivers : a representative sample from the salmon rivers located in western Europe and under the influence of the Gulf Stream.
Biological Reference Points : spawning target S^*, maximum sustainable exploitation rate h^*.

River	Country	Latitude (°N)	Riverine wetted area accessible to salmon (m^2)	Number of SR observations (years)
Nivelle	France	43	320995	12
Oir	France	48.5	48000	14
Frome	England	50.5	876420	12
Dee	England	53	6170000	9
Burrishoole	Ireland	54	155000	12
Lune	England	54.5	4230000	7
Bush	N. Ireland	55	845500	13
Mourne	N. Ireland	55	10360560	13
Faughan	N. Ireland	55	882380	11
Girnock Burn	Scotland	57	58764	12
North Esk	Scotland	57	2100000	6
Laerdalselva	Norway	61	704000	8
Ellidaar	Iceland	64	199711	10


Laerdalselva
Ellidaar
Girnock B.
North Esk
Bush
Mourne
Faughan
Burrishoole
Lune
Dee
Frome
Oir
Nivelle
N
W
E
S
67
62
57
52
47
42
Latitude
-25
-20
-15
-10
-5
0
5
10
15
Longitude

Independent vs hierarchical SR analyses



The Atlantic salmon SR series on the 13 index rivers and fitted SR relationships. S (x-axis) and R (y-axis) are the stock and recruitment variables after standardization for river size expressed in eggs per m^2 of riverine wetted area accessible to salmon. The SR Ricker curves are graphed for two model configurations, the model assuming independence between rivers (thin line) and the hierarchical model (bold line). SR curves are graphed with parameters (S^*, h^*) set at the median of their marginal posterior distributions.

Questions

- How is the SR information transferred from the monitored data-rich rivers to set Biological Reference Points for other sparse-data salmon rivers, while accounting for the major sources of uncertainty ?

- How can the joint analysis of the SR relationship for the 13 index rivers be used to forecast biological reference points for a new river without any SR data but for which relevant covariates are available ?

Likelihood = LogNormal + Ricker

For river $k = 1, ..., 13$, one relates the recruitment $R_{k,t}$ rescaled in $eggs/m^2$ of the cohort born in year t to the associated spawning stock $S_{k,t}$:

$$\begin{cases} log(R_{k,t}) = h^*_k + log(\frac{S_{k,t}}{1 - h^*_k}) - \frac{h^*_k}{S^*_k} S_{k,t} + \varepsilon_{k,t} \\ \varepsilon_{k,t} \overset{iid}{\sim} Normal(0, \sigma^2_k) \end{cases} \quad (1)$$

where σ_k is the standard deviation of the Normal distribution of $log(R_{k,t})$, S^*_k and h^*_k are respectively the stock which are necessary to guarantee an optimal sustainable exploitation and the associated optimal exploitation rate for the river k.

Priors

Poor informative priors, independently for each river k :

$$\begin{cases} h_k^* \sim Beta(1,1) \\ S_k^* \sim Uniforn(0,200) \end{cases}$$

More refined elicitation of S^* : $\mu_{S^*} = 40$ eggs per m^2 with $\sigma_{S^*} = 40$ so that $CV_{S^*} = \sigma_{S^*}/\mu_{S^*} = 1$
Gamma pdf with shape parameter a and scale parameter b such that $\mu_{S^*} = \frac{a}{b}$, $CV_{S^*} = \frac{1}{\sqrt{a}}$ and constrained to the range $]0,200]$:

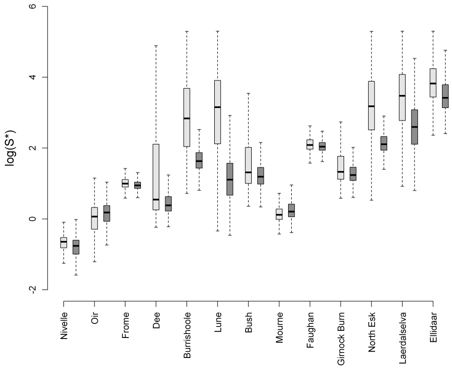
$$\begin{cases} \mu_{S^*} = 40\ eggs/m^2;\ CV_{S^*} = 1 \\ a = \dfrac{1}{CV_{S^*}^2};\ b = \dfrac{1}{\mu_{S^*} \times CV_{S^*}^2} \\ S_k^* \sim Gamma(a,b)1_{S_k^* < 200} \end{cases}$$

σ_k^2 constant across all the rivers in the study, and a Gamma prior distribution was assigned on the precision :

$$\begin{cases} \forall k,\ \sigma_k = \sigma \\ \sigma^{-2} \sim Gamma(p,q) \end{cases}$$

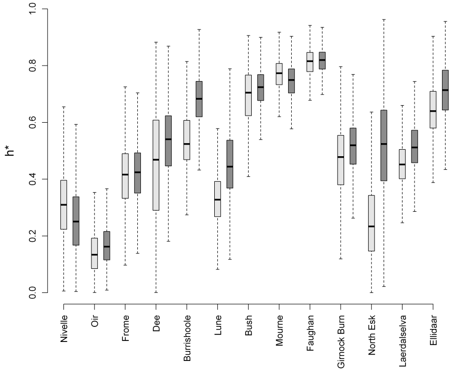
Diffuse $Gamma(p,q)$ prior for the precision $\sigma^{-2}$by letting p and q being very small.

Posterior pdf of Log(Sustainable stock)



Marginal posterior distribution of $log(S^*)$ (in eggs per m^2) for the 13 index rivers obtained under two model configurations : the model assuming independence between rivers (light gray) and the hierarchical model (dark gray).

Exploitation rate h posterior pdf

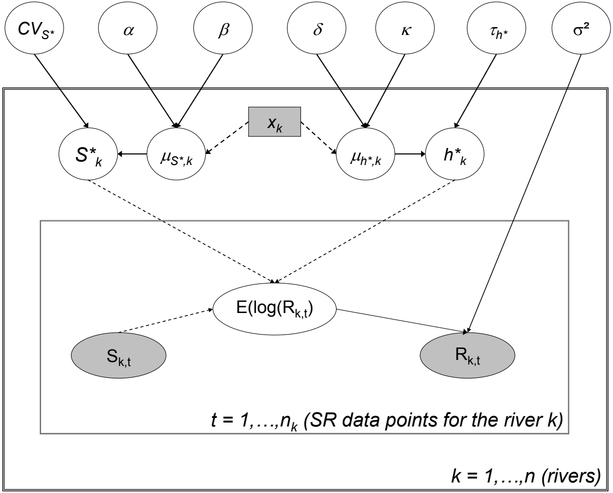


Marginal posterior distribution of h^* for the 13 index rivers obtained under two model configurations : the model assuming independence between rivers (light gray) and the hierarchical model (dark gray).

Posterior inferences on the model assuming independence between rivers

- All distributions exhibit quite heavy tails.
- Depending on the river, the number of observations, and on the contrast between the S values in the observation sample, uncertainty range in the posterior inferences may differ from several orders of magnitude.
- The boxplots of the parameters' posterior pdfs reveal an increasing latitudinal gradient in the S_k^*'s.

Posterior inferences on the hierarchical model with partial exchangeability between rivers (Latitude as a covariate)



Posterior inferences on the hierarchical model with partial exchangeability between rivers (Latitude as a covariate)

S_k is related to latitude of river k denoted x_k :

$$\begin{cases} log(\mu_{S_k^*}) = \alpha \times x_k + \beta \\ \alpha \sim Uniform(-5,5) \\ \beta \sim Uniform(-50,50) \end{cases}$$

Parameters a_k and b_k may depend upon the latitude x_k :

$$\begin{cases} CV_{S^*} \sim Uniform(0,20) \\ a_k = \dfrac{1}{CV_{S^*}^2} \\ b_k = \dfrac{1}{\mu_{S_k^*} \times CV_{S^*}^2} \\ S_k^* \sim Gamma(a_k, b_k) 1_{S_k^* < 200} \end{cases}$$

Posterior inferences on the hierarchical model with partial exchangeability between rivers (Latitude as a covariate)

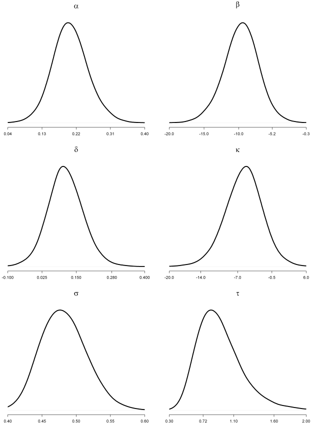
Residual degree of similarity between rivers once the latitude gradient is accounted for :

$$\begin{cases} logit(\mu_{h_k^*}) = \delta \times x_k + \kappa \\ \delta \sim Uniform(-5, 5) \\ \kappa \sim Uniform(-50, 50) \end{cases}$$

Diffuse prior set on precision :

$$\begin{cases} logit(h_k^*) \sim Normal(logit(\mu_{h_k^*}), \tau^2) \\ \tau_{h^*}^{-2} \sim Gamma(0.001, 0.001) \end{cases}$$

Posterior inferences on the hierarchical model with partial exchangeability between rivers



Marginal posterior probability shapes of the parameters α, β, δ, κ, σ and τ from the hierarchical model.

Posterior inferences on the hierarchical model with partial exchangeability between rivers

The covariate latitude offers a good statistical explanation of positive variations between rivers in both h^* and S^*. Posterior pdf of S^* and h^* for the monitored 13 rivers reveal :

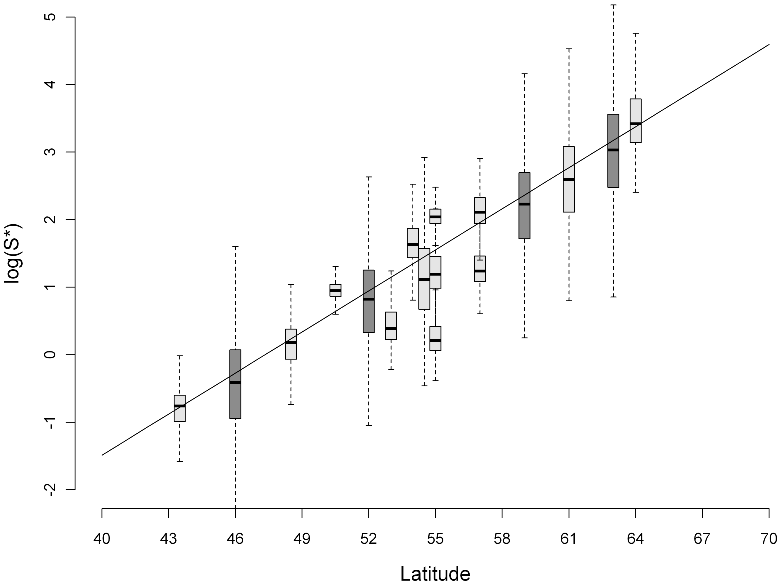
- Considerable within-river uncertainty in some cases despite SR data being available (*e.g.*, the Lune R. and the Laerdalselva R.) ;
- Significant variations among rivers, even within a relatively narrow latitudinal range (*e.g.*, the Bush R., the Mourne R. and the Faughan R., all located in Northern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

Posterior inferences on the hierarchical model between rivers

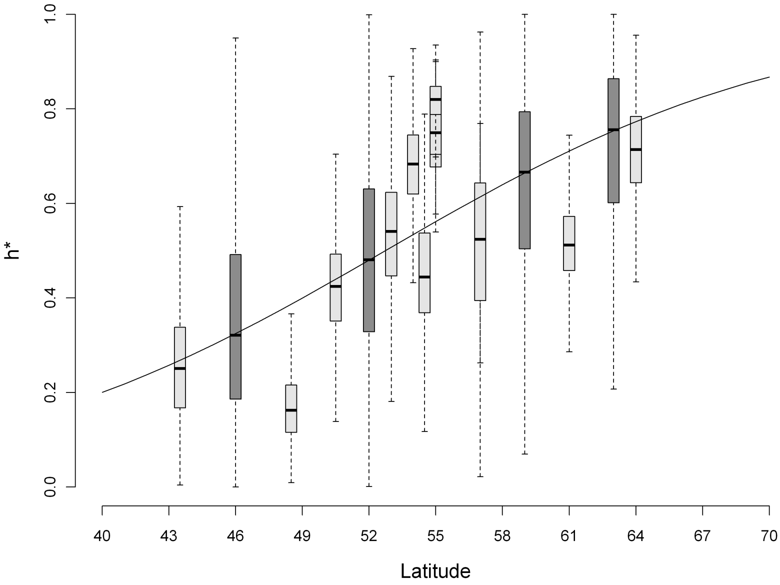
h^*_{new} and S^*_{new} at various latitudes covering the salmon distribution range in the northeast Atlantic area (46°, 52°, 59° and 63° north) .

- Moving north, salmon stocks can sustain higher exploitation rates h^*, can produce higher recruitment at MSY ($R^* = \frac{S^*}{1-h^*}$), but at the same time should be set at higher conservation limits S^*.
- Great remaining uncertainty in the spawning stock, the recruitment and the exploitation rates at MSY for a sparse data river
- important residual variations stem from other explanatory covariates than the riverine wetted area and the latitude.

Posterior inferences on the hierarchical model



Posterior inferences on the hierarchical model



References

J.M. Bernardo and A.F.M. Smith. *Bayesian Theory*.
Wiley & Sons, Series in Probability and Mathematical Statistics, London, 1994.

D. Collett. *Modelling Binary Data*.
Chapman & Hall, 2nd edition, 2003.

N.A.C. Cressie, C.A. Calder, J.S. Clark, J.M. Ver Hoeff, and C.K. Wikle.
Accounting for uncertainty in ecological analysis : the strenghts and limitations of hierarchical statistical modelling.
Ecological Applications, 19(3) :553–570, 2009.

W.W. Crozier, E.C.E. Potter, E. Prévost, P.-J. Schön, and O. Maoiléidigh.
A coordinated approach towards the development of a scientific basis for management of wild Atlantic salmon in the North-East Atlantic.
Technical report, Concerted Action QLK5-CT1999e01546 (SALMODEL), 2003.

B. de Finetti.
La Prévision : ses Lois Logiques, ses Sources Subjectives.
Institut Henri Poincaré, Paris, 1937.

D. Dey, S. K. Ghosh, and B.K Mallick.
Generalized Linear Models : A Bayesian Perspective.
CRC Press, Boca Raton, FL, 2000.

B. Efron and C. Morris.
Data analysis using Stein's estimator and its generalizations.
Journal of the American Statistical Association, pages 311–319, 2004.

References

L. Fahrmeir and G. Tutz.
Multivariate Statistical Modelling Based on Generalized Linear models.
Springer-Verlag, New York, NY, 1994.

E. Hewitt and L.J. Savage.
Symmetric measures on Cartesian products.
Transactions of the American Mathematical Society, 80 :470–501, 1955.

W.A. Link, E. Cam, J.D. Nichols, and E.G. Cooch.
Of bugs and birds : Markov chain Monte Carlo for hierarchical modeling in wildlife research.
Journal of Wildlife Management, 66(2) :277–291, 2002.

C.E. McCulloch, S.R. Searle, and J.M. Neuhaus.
Generalized, Linear and Mixed Models.
Wiley, New York, NY, 2008.

E. Prévost, G. Chaput, and (Ed.).
Stock, Recruitment and Reference Points Assessment and Management of Atlantic salmon.
INRA Editions, Paris, 2001.

E. Prévost, E. Parent, W. Crozier, I. Davidson, J. Dumas, G. Gudbergsson, K. Hindar, P. McGinnity, J. MacLean, and L.M. Sattem.
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, 2003.

References

C.R. Rao and H. Toutenburg.
Linear Models : Least Squares and Alternatives.
Springer, New York, NY, 1999.

J.T. Schnute and A.R. Kronlund.
A management oriented approach to stock recruitment analysis.
Canadian Journal of Fisheries and Aquatic Sciences, 53 :1281–1293, 1996.

D. Sorensen and D. Gianola.
Likelihood, Bayesian, and MCMC Methods in Quantitative Genetics.
Springer, New York, NY, 2002.

C. Stein.
Inadmissibility of the usual estimator for the mean of a multivariate distribution.
Proceedings of the Third Berkeley Symposium on Mathematical and Statistical Probability, 1 :197–206, 1956.

G. Verbeke and G. Molenberghs.
Linear Mixed Models for Longitudinal Data.
Springer, New York, NY, 2000.

C.K. Wikle.
Hierarchical Bayesian models for predicting the spread of ecological processes.
Ecology, 84(6) :1382–1394, 2003.