



AGROCAMPUS OUEST

65 rue de Saint Briec
CS 84215
35042 Rennes Cedex
Tél : 02 23 48 55 00

INRA/UMR ESE

65 rue de Saint Briec
CS 84215
35042 Rennes Cedex

Mémoire de Fin d'Etudes

Pour l'obtention du Diplôme d'Agronomie Approfondie (DAA)

Spécialisation Halieutique

**Spatio-temporal variability of Atlantic salmon
juvenile abundance in a Canadian stream:
A Bayesian modeling approach.**

Présenté par : Mlle BOULENGER Clarisse

Soutenu le : 16/09/2011



AGROCAMPUS OUEST

65 rue de Saint Briec
CS 84215
35042 Rennes Cedex
Tél : 02 23 48 55 00

INRA/UMR ESE

65 rue de Saint Briec
CS 84215
35042 Rennes Cedex

Mémoire de Fin d'Etudes

Diplôme d'ingénieur agronome

Spécialisation ou spécialité : Halieutique

Option: REA

**Spatio-temporal variability of Atlantic salmon juvenile abundance in a
Canadian stream:
A Bayesian Modeling Approach**

Par : Mlle BOULENGER Clarisse

Soutenu le : 16/09/2011

Devant le jury :

M. Gascuel Didier, Agrocampus Ouest

M. Plantegenet Manuel, Agrocampus Ouest

Mme. Trenkel Verena, Ifremer Nantes

M. Jean-Marc Roussel, INRA Rennes

Maître de stage : Etienne Rivot

Co-encadrant : Richard Cunjak (Absent lors de la soutenance)

Bon pour dépôt

Autorisation de diffusion externe du mémoire : Oui Non

"Les analyses et les conclusions de ce travail d'étudiant n'engagent
que la responsabilité de son auteur et non celle d'AGROCAMPUS OUEST".

Diffusion du mémoire

A remplir par l'auteur avec le maître de stage.

Aucune confidentialité ne sera prise en compte si la durée n'en est pas précisée.

Préciser les limites de la confidentialité ⁽¹⁾ :

Mémoire de fin d'études

Consultable sur place : oui non
Reproduction autorisée : oui non
Prêt autorisé : oui non
Confidentialité absolue : oui non
(ni consultation, ni prêt)

Si oui 1 an 5 ans 10 ans

Résumé diffusable : oui non

Diffusion de la version numérique : oui non

Si oui, l'auteur complète l'autorisation suivante :

Je soussigné(e) Christine Boulenger, propriétaire des droits de reproduction dudit résumé, autorise toutes les sources bibliographiques à le signaler et le publier.

Date : 19/09/11

Signature : 

Rennes, le 19/09/11

Le Maître de stage⁽²⁾,



L'auteur,



L'Enseignant responsable⁽²⁾,



Didier GASCUEL
Professeur Directeur du Pôle halieutique
AGROCAMPUS OUEST

(1) L'administration, les enseignants et les différents services de documentation d'AGROCAMPUS OUEST s'engagent à respecter cette confidentialité.

(2) Signature et cachet de l'organisme.

Acknowledgements

Je souhaite tout d'abord remercier M. Etienne Rivot pour son encadrement de qualité qui m'a permis d'acquérir des connaissances solides en modélisation bayésienne aussi bien dans les principes que dans l'utilisation des outils permettant sa mise en œuvre. Je le remercie également pour sa disponibilité, ses conseils, ses corrections toujours expliquées et son écoute aussi bien pour la réalisation de ce stage que pour mon choix de thèse.

Je remercie également Richard Cunjak, Etienne Prevost et Jean-Marc Roussel pour leur co-encadrement. Merci à Rick pour m'avoir permis de travailler sur ce jeu de données. Je le remercie aussi ainsi que Jean-Marc pour leur accompagnement sur les questions écologiques de ce stage et pour leur patience et leur écoute lors de la phase de modélisation. Merci à Etienne Prevost pour son accueil lors de ma visite à St Pée sur Nivelles et pour sa disponibilité et ses conseils lors de ce stage.

Je tiens à remercier l'ensemble du pôle Halieutique d'Agrocampus Ouest auquel j'étais rattachée, pour leur accueil, leur gentillesse, leur disponibilité et leur aide précieuse. Merci aux autres stagiaires du couloir, Axel, Gwen, Caro, Cha, Gaël et Ronan pour leurs sourires et leur soutien.

Merci à l'homme du fond d'écran qui m'a permis de m'évader à chaque ouverture de mon ordinateur. Merci également à Clémentine pour nos discussions salvatrices.

Enfin, je remercie mes proches qui m'ont aidé tout au long de mon parcours et plus particulièrement Isabelle et Cyril qui même pendant les weekends end me font travailler en allant visiter Eog.

Table of contents

INTRODUCTION	1
MATERIAL AND METHODS	5
STUDY SITE AND DATA SET	5
<i>Catamaran brook</i>	5
<i>Sampling sheme and data</i>	6
HIERARCHICAL BAYESIAN MODEL	7
<i>Capture model for one site (a, h, y, r, k)</i>	8
<i>Hierarchical structure between sites</i>	9
Probability of capture.....	9
<i>A mixture hierarchical model for the Density</i>	9
A Gamma distribution for the density.....	9
A mixture model to capture the variability in accessibility.....	9
Summing up the model specification in a Directed Acyclic Graph.	11
<i>Prior specification for free parameters</i>	11
<i>Extrapolation of the number of fish to the whole Catamaran watershed</i>	12
<i>Posterior Checking</i>	13
<i>Bayesian Computations</i>	13
RESULTS	14
POSTERIOR DISTRIBUTION OF KEY PARAMETERS AND LATENT VARIABLES, AND CHECKING OF KEY MODELLING	
HYPOTHESES	14
<i>Capture probability</i>	14
<i>Correlation between p and d ?</i>	15
<i>Effects on density</i>	15
Mixture model.....	15
Interaction Year, Age, Reach	17
Habitat preferences.....	18
<i>Posterior checking</i>	18
ABUNDANCE OF JUVENILES AND POPULATION DYNAMICS AT THE SCALE OF THE CATAMARAN BROOK	20
<i>Total abundance of 0+, 1+ and 2+ juveniles</i>	20
<i>First insight on the Stock-Recruitment relationship</i>	20
<i>Ratio N1+/N0+ and ratio N2+/N1+</i>	22
FUNCTIONAL ROLE OF THE DIFFERENT REACHES WITH REGARDS TO THE POPULATION DYNAMICS	23
<i>Relative contribution of the reaches to the total abundance</i>	23
<i>Relationship between accessibility and environmental covariates</i>	25
<i>Evidence of an intense migration between reaches</i>	26
DISCUSSION	28
MODEL.....	28
POPULATION DYNAMICS	29
PROSPECTS.....	31
REFERENCES	33

List of Appendix

APPENDIX I : PROBABILITY DISTRIBUTIONS AND FUNCTIONS USED IN THE MODEL 37
APPENDIX II: WINBUGS[®] MODEL 38

List of illustrations

Figures

FIG 1: DISTRIBUTION AREA OF ATLANTIC SALMON (<i>SALMO SALAR</i>) IN NORTH ATLANTIC. ARROWS REPRESENT THE MIGRATION TO THE SEA OF THE DIFFERENT POPULATIONS OF ATLANTIC SALMON (SOURCE : ATLANTIC SALMON FEDERATION – WWW.ASF.CA/DOCS/UPLOADS/NORTHATLANTICUP-L.JPG).....	1
FIG 2 : LIFE CYCLE OF ATLANTIC SALMON (SOURCE : WITH COURTESY OF THE ATLANTIC SALMON TRUST AND ROBIN ADE – WWW.NASCO.INT/ATLANTICSALMON.HTML).....	2
FIG 3: SITUATION OF CATAMARAN BROOK IN NEW BRUNSWICK, CANADA. THE BOTTOM PANEL GIVES THE POSITIONS OF THE 5 REACHES SAMPLED IN THE BROOK.....	6
FIG 4 : SCHEME OF THE HIERARCHICAL BAYESIAN MODEL FOR INTEGRATING SUCESSIVE REMOVAL DATA. THE FIRST LAYER (BLUE) REPRESENTS THE NUMBER OF FISH CAUGH IN THE DIFFERENT SITES EACH YEAR. THE SECOND LAYER (YELLOW) SYMBOLIZES THE ESTIMATED DENSITY FOR EACH SITE. THE THIRD LAYER CORRESPONDS TO THE PARAMETERS AND EFFECTS ESTIMATED THANKS TO THE DATA. BLACK ARROWS REPRESENT THE MODELING PROCESS AND THE STATISTICAL INFERENCE FOLLOWS THE REVERSE WAY.	8
FIG 5 : SIMPLIFIED DIRECTED ACYCLIC GRAPH (DAG) FOR THE MODEL	11
FIG 6 : PROBABILITY OF CAPTURE OF ALL SAMPLING UNITS DRAWN AS A FUNCTION OF A)AGE, B) REACH AND AGE, C) HABITAT AND AGE, D) YEAR AND AGE. BOXPLOTS REPRESENTS THE BETWEEN UNITS VARIABILITY OF POINT ESTIMATES (POSTERIOR MEDIAN) OF P. RED, GREEN AND BLUE REPRESENT RESPECTIVELY 0+,1+ AND 2+	15
FIG 7 : PAIRWISE MCMC PLOTS SHOWING THE SHAPE OF THE JOINT POSTERIOR DISTRIBUTION OF $\mathbf{p(a, h, y, r, k)}$, $\mathbf{d(a, h, y, r, k)}$ FOR 2 EXAMPLES : A) LOWER, 2006, 2+, BEDROCK ; B) MIDDLE, 2002, 2+, RUN	16
FIG 8 : TIME SERIES OF THE POSTERIOR MEAN OF THE BERNOULLI VARIABLE $\mathbf{B(a, y, r)}$ INTERPRETED AS AN INDICATOR OF THE PRESENCE OF FISH OF THE THREE AGE CLASSES IN EACH REACH BETWEEN 1990-2010. THE AVERAGE POSTERIOR PROBABILITY THAT $B=1$ FOR THE 3 AGES CLASSES ARE RESPECTIVELY (0.97, 0.97, .0.97), (0.97, 0.97, 0.97), (0.56, 0.90, 0.94), (0.65, 0.86, 0.51), (0.04, 0.21, 0.07), (0.04, 0.12, 0.13) RESPECTIVELY FOR LOWER, GORGE, MIDDLE, OTTER BROOK, UPPER AND TRIBUTARY.	16
FIG 9 : POSTERIOR DISTRIBUTION OF $\beta d_1(a, y, r)$ PER AGE CLASS AND REACH, THE HORIZONTAL LINES REPRESENT THE MEAN OF βd_1 FOR EACH AGE CLASS AND REACH DURING THE 21 YEARS (REPRESENTED IN THE X AXIS). BOXPLOT IN WHITE FONT ARE THE POSTERIOR PREDICTIVE OF THE βd_1 'S FOR THE COMBINATION OF FACTORS WHERE NO ELECTROFISHING DATA WERE AVAILABLE	17
FIG 10 : POSTERIOR DISTRIBUTION OF HABITAT EFFECTS PER AGE CLASS $\mathbf{ad(a, h)}$. THE HORIZONTAL LINE REFLECT THE ABOVE AVERAGE OF $\mathbf{ad(a, h)}$ (EQUAL TO 0 BY CONSTRUCTION).....	18
FIG 11 : COMPARISON BETWEEN $\mathbf{d(a, h, y, r, k)}$ AND $\mathbf{d'(a, h, y, r, k)}$ FOR ALL SAMPLING SITES. THE RED LINE REPRESENTS THE VALUE 0.....	19
FIG 12 : COMPARISON BETWEEN $\mathbf{D(A,H,Y,R,K)}$ AND $\mathbf{D'(A,H,Y,R,K)}$ FOR A) THE UPPER REACH AND B) THE TRIBUTARY REACH. THE RED LINE REPRESENTS THE VALUE 0	19
FIG 13 : TIME SERIES OF ESTIMATED ABUNDANCE OF 0+, 1+ AND 2+ SALMON JUVENILES IN CATAMARAN BROOK, 1990-2010. 1+ AND 2+ HAVE BEEN ONE (RESP. TWO) YEAR LAGGED TO MATCH WITH THE 0+ FISH, SO THAT FISH FROM THE SAME COHORT ARE ALIGNED ON THE SAME YEAR. FOR EACH AGE CLASS, THE THICK LINE REPRESENTS THE POSTERIOR MEDIAN AND THE SHADED AREA IS THE POSTERIOR UNCERTAINTY (97.5% CREDIBILITY INTERVAL)	20
FIG 14 : RELATIONSHIP BETWEEN THE DENSITY OF 0+ AND A) THE NUMBER OF SPWNERS MIGRATING UP IN THE CATAMARAN BROOK (ONE YEAR LAGGED) (THE SIZE OF THE CIRCLES IS PROPORTIONAL TO THE AVERAGE WATER DISCHARGE BETWEEN OCTOBER AND NOVEMBER); B) THE AVERAGE WATER FLOW (OCT-NOV; ONE YEAR LAGGED) (THE SIZE OF THE CIRCLES IS PROPORTIONAL TO THE NUMBER OF SPAWNERS).....	21
FIG 15 : RELATIONSHIP BETWEEN THE NUMBER OF SPAWNERS MIGRATING UP IN THE CATAMARRAN BROOK AND THE AVERAGE WATER DISCHARGE (OCTOBER-NOVEMBER).....	21
FIG 16: RELATIONSHIP BETWEEN THE RATIO (NUMBER OF 0+/NUMBER OF SPAWNERS) AND A) THE NUMBER OF SPAWNERS ; B) THE AVERAGE WATER DISCHARGE (OCTOBER-NOVEMBER).....	22
FIG 17 : TIME SERIES OF POSTERIOR ESTIMATES OF A) THE RATIO $\log(N_{1+}/N_{0+})$. 1+ FISH HAVE BEEN ONE YEAR LAGGED TO MATCH WITH THE 0+ FISH B) THE RATIO $\log(N_{2+}/N_{1+})$. 2+ FISH HAVE BEEN ONE YEAR LAGGED TO MATCH WITH THE 1+ FISH.. THICK LINE: POSTERIOR MEDIAN; SHADED AREA: 97.5% BAYESIAN POSTERIOR CREDIBILITY INTERVAL.	23

FIG 18 : TIME SERIES OF THE RELATIVE CONTRIBUTION OF THE 4 MAIN REACHES TO THE OVERALL ABUNDANCE FOR THE THREE AGE-CLASSES (0+, 1+ AND 2+) OF JUVENILE SALMON IN THE CATAMARAN BROOK.....	24
FIG 19 : RELATIVE CONTRIBUTION OF THE 4 MAIN REACHES TO THE OVERALL ABUNDANCE FOR THE THREE AGE-CLASSES (0+, 1+ AND 2+) OF JUVENILE SALMON AS DEMONSTRATED FOR FOUR COHORTS 1993, 1996, 1997 AND 2004 (YEAR OF BIRTH OF 0+ JUVENILES).	24
FIG 20 : RELATION BETWEEN THE RELATION CONTRIBUTION OF THE MIDDLE REACH AND THE ACCESSIBILITY (B) OF THIS SECTOR A) 0+, B) 1+, C) 2+	25
FIG 21 : POSTERIOR MEDIAN OF THE INDICATOR VARIABLE B (INTERPRETED AS AN INDICATOR OF THE ACCESSIBILITY) FOR 0+ JUVENILES AS A FUNCTION OF THE AVERAGE DISCHARGE (OCT-NOV) OF A) LOWER REACH, B) MIDDLE REACH.	26
FIG 22 : TIME SERIES OF THE RATIO A) (NUMBER OF 1+ FISH / NUMBER OF 0+ FISH) (IN LOG-SCALE); B) (NUMBER OF 2+ FISH / NUMBER OF 1+ FISH) IN THE MIDDLE REACH. THE 1+ HAVE BEEN ONE YEAR LAGGED WITH RESPECT TO 0+ FISH, AND THE 2+ FISH HAVE BEEN ONE YEAR LAGGED WITH RESPECT TO THE 1+ FISH.	27

Tables

TABLE 1: CHARACTERISTICS OF THE FIVE HABITAT TYPES PRESENT IN CATAMARAN BROOK (SOURCE : CUNJAK ET AL., 1993, GIRARD ET AL., 2004).	5
TABLE 2 : SURFACE AREA (M ²) AND RANGE OF PERCENTAGE OF SURFACE AREA SAMPLED (DURING THE 21 YEARS) PER REACH AND HABITAT TYPE.....	7
TABLE 3: PRIOR DISTRIBUTION OF FREE PARAMETERS	12
TABLE 4 : MAIN STATISTICS OF THE POSTERIOR DISTRIBUTION OF PARAMETERS FOR THE HIERARCHICAL STRUCTURE ON CAPTURE PROBABILITY. $M_P = \text{LOGIT}(M_Z)$	14

Introduction

Understanding the ecology of fish populations in interaction with their ecosystem is a major challenge of aquatic and fisheries ecology and is a prerequisite for managing populations and ecosystems.

Because of their high life-history variability and their sensibility to the different pressures throughout their life cycle and at different spatial scales (marine and freshwater phase), diadromous fish (*e.g.* salmon, shads, lamprey, eels ...) have received a considerable attention from population ecologists.

Atlantic salmon (*Salmo salar*) occurs naturally along both East and West coasts of the North Atlantic Ocean (Fig.1) (Mills, 1989; Shearer, 1992; Klemetsen et al., 2003). Many populations of Atlantic salmon are now extinct or seriously endangered (Webb, 2007; Limburg and Waldman, 2010). The fragmentation of habitat (due for example to dams), the degradation of water quality due to human activities and overfishing are the main causes of this rarefaction (Parrish et al., 1998; Limburg and Waldman, 2010). Some *A. salmon* populations are now listed as endangered, in the USA (Maine) (US Endangered Species Act), in Canada (inner Bay of Fundy, Nova Scotia, New Brunswick) (Committee on the Status of Endangered Wildlife), or in Europe (registration in the appendix II of “EU Habitat Framework” which implies the definition of conservation areas for this species). Canada has the most important part of Eastern Atlantic wild salmon populations.

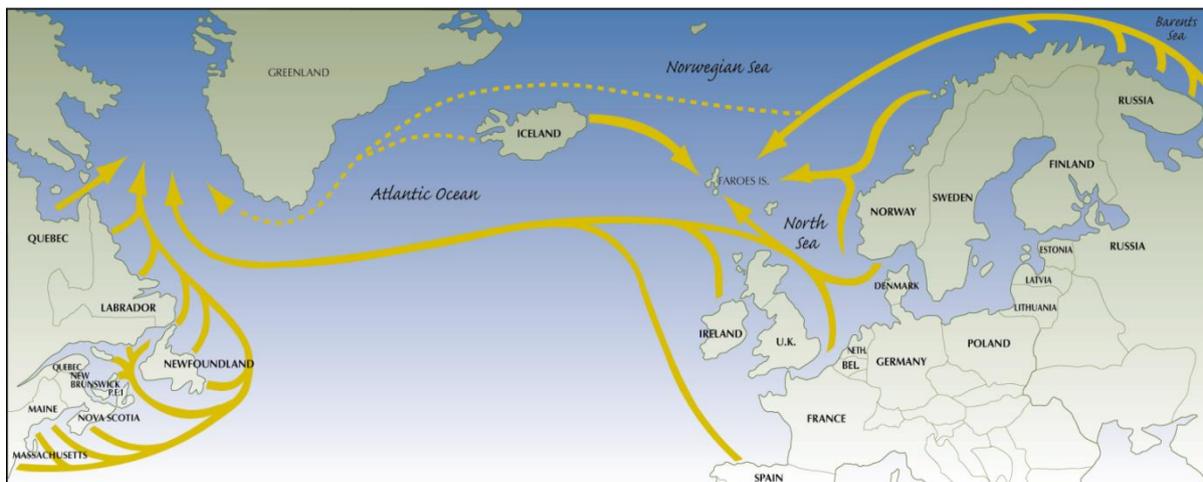


Fig 1: Distribution area of Atlantic salmon (*Salmo salar*) in North Atlantic. Arrows represent the migration to the sea of the different populations of Atlantic salmon (Source : Atlantic Salmon Federation – www.asf.ca/docs/uploads/northatlanticup-1.jpg)

The life cycle of *A. salmon* is divided between the river and the ocean (Fig. 2). In Canada, adults migrate upstream to spawn in late autumn (October- November) (Mitchell and Cunjak, 2007a). Eggs then Juveniles (0+, 1+ and 2+) stay in freshwater until smoltification, an important morphological and physiological transformation preparing them for sea life. Smolts leave the river and migrate to the Ocean where they will live between one and three year before returning to the freshwater tributary in which they were born. Initial movement of fish from the sea to a location close to their ultimate spawning destination generally occurs between June and August (Mitchell and Cunjak, 2007a).



Fig 2 : Life cycle of Atlantic salmon (Source : With courtesy of the Atlantic Salmon Trust and Robin Ade – www.nasco.int/atlanticsalmon.html)

Understanding the factors controlling the spatio-temporal dynamics of freshwater juvenile (i.e. parr) abundance in the wild is of primary interest for population ecology. It is critical for the analysis of stock (i.e., egg deposition by spawners) and recruitment (i.e., juvenile production of the subsequent generation) relationships. The transition from eggs to 0+ juvenile is a major bottleneck for Atlantic salmon populations (Elliott 2001; Milner *et al.*, 2003). From a management perspective of salmon stocks, the freshwater juvenile is the earliest, the most widely used, and often the only development stage that can be monitored prior to and independently from fisheries.

The abundance of A. salmon juveniles in freshwater is highly spatially structured within a watershed. Numerous factors control the spatial distribution of juveniles. (1) First, juvenile abundance depend upon the distribution of spawners (and associated redds) in the watershed during the spawning season. This can be controlled by obstacles (natural or human made) in interaction with water discharge (Mitchell and Cunjak, 2007b; Taylor *et al.*, 2010). Preferendum for spawning habitat also conditions the initial distribution of 0+ juveniles; (2) Second, density dependent and independent factors can control the eggs \rightarrow 0+, 0+ \rightarrow 1+ and older survival rates. These factors are also susceptible to be highly variable both in space and time (Tezlaff *et al.*, 2005); (3) Third, density dependent and independent factors control the migration of juveniles within the watershed. In particular, upstream and downstream migrations have been reported between 0+ and 1+ stages, and between 1+ and 2+ stages (Steingrímsson *et al.*, 2003).

Hence, analyzing the spatio-temporal distribution of A. salmon juveniles is a prerequisite for understanding the functional role of different sectors within a watershed and of different habitat type within each sector with regards to the recruitment dynamics and to the whole population dynamics.

Successive removal by electrofishing is the most commonly used method for deriving estimates of abundance of riverine fish such as salmonids (Bohlin *et al.*, 1990). First, a

number of sites (typically characterized by different position in the watershed and different habitat type) are selected from the river stretch of interest and then the fish are sampled from each site by the successive removal method (Hankin 1984; Bohlin *et al.*, 1990). In such kind of survey, row data then consist in a series of number of fish caught during the successive removals pass in different sites. Interpreting these field surveys require integrating these row catches data within statistical models aiming at (i) Estimating population size (or density) at the sampling site level; (ii) Integrating these estimations obtained at the site level to quantify the different factors (time, sectors, habitat type ...) structuring the variability of the density; (iii) Extrapolating these results to quantify juvenile abundance in space and time at the scale of the whole river stretch.

Hierarchical Bayesian Models (HBM) provide a useful framework for such statistical analyzes (Wyatt, 2002; Wyatt *et al.*, 2003; Rivot *et al.*, 2008; Dauphin *et al.*, 2009; Ruiz and Laplanche, 2010; Brun *et al.*, 2011). They offer a consistent framework for the analysis of large and heterogeneous data set characterized by data rich and data poor statistical units and multiple dependencies between variables. They can handle large and complex models in term of the number of the structure of dependency between the variables. The probabilistic structure of hierarchical models helps to jointly treat data rich and data poor statistical units by organizing the transfer of information from data rich to data poor units (Rivot and Prévost, 2002). HBM also offer a consistent probabilistic framework to explicitly consider the different sources of uncertainty in estimations and extrapolations (predictions).

In this work, a HBM was built to analyse the spatio-temporal variability of the abundance of three age classes of *A. salmon* juveniles (0+, 1+ and 2+) in the Catamaran brook (New Brunswick, Canada) during 21 years between 1990 and 2010. The data set analyzed in this work is highly informative. Each year between 1990 and 2010, between 9 and 31 sites (characterized by a position in the watershed and a habitat type) were sampled for 0+, 1+ and 2+ juveniles density through electrofishing with 3 to 5 successive removal pass. Environmental variables such as water discharge are also available.

An original statistical HBM was built to analyze these data. In particular, by analogy of zero-inflated models (Martin *et al.*, 2005; Sileshi *et al.*, 2009; Ancelet *et al.*, 2010, Calama *et al.*, 2011; Vaudor *et al.*, 2011), the HBM integrates a mixture component in the distribution of the fish density that can handle both very low and very high fish density, and this was analyzed in relation with environmental factors (*e.g.* water discharge) that may condition accessibility to some river reaches.

The model was designed to explore interesting ecological questions. In particular, the objectives were:

- (1) To estimate the abundance of the three juvenile age classes (0+, 1+ and 2+) at the scale of the Catamaran brook between 1990 and 2010;
- (2) To analyze the stock-recruitment dynamics (spawners \rightarrow 0+) and the dynamics of 0+ \rightarrow + and 1+ \rightarrow 2+ survival rates at the scale of the whole river stretch, and propose a first exploration of the relationship with environmental variables such as water discharge;
- (3) To quantify the functional role of the different reaches with regards to the contribution to the juveniles recruitment (3 age classes) at the scale of the whole Catamaran watershed;
- (4) Explore the relationship between the spatio-temporal distribution of juveniles and the environmental factors (*e.g.* water discharge) controlling the accessibility of different reaches to spawners.
- (5) Explore if a link between the patterns of colonization of different reaches in the watershed and the recruitment dynamics exists.

The report is organized as follows. In the first part of the Material and Methods section, the study site and the data set are detailed. Then, the HBM model is described. Results first propose some diagnostics to evaluate the modeling hypotheses. The second part of the results section concentrates on estimates of juvenile abundance at the scale of the whole watershed and on stock-recruitment relationship. Results concerning the contribution of the different reaches within the watershed are given in the third part. In the discussion, the originality of the HBM approach are outlined, and the ecological meaning of our results are discussed. The report concludes with some interesting perspectives of this work.

Material and methods

Study site and data set

Catamaran brook

The Catamaran brook (46°52.7'N, 66°06.0'W) is a third-order tributary of the Little Southwest Miramichi River in central New Brunswick, Canada (Fig. 3). The Catamaran brook is a remarkable study area with a long term multidisciplinary research program (Cunjak et al., 1993), with numerous studies about dynamic and behaviour of juveniles (Cunjak and Therrien, 1998; Steingrimsson et al.; 2003, Girard et al.; 2004, Imre et al., 2005, Breau et al.; 2007).

The drainage area is about 52 km² and the main watercourse is 20.5 km with a mean slope of 1.3% (Cunjak et al., 1990). Atlantic salmon is the most common fish species in this stream but the distribution is not homogenous due to the presence of beaver dams and waterfall that limits the colonisation of the upstream part of the system (Cunjak and Therrien, 1998). Adult Atlantic salmon return to spawn in the river between late September and early November (Cunjak et al., 1993). Salmon juveniles remain in the stream for 2–3 years before migrating to sea (Randall, 1982; Cunjak et al., 1993). Five stream sections were chosen (Fig. 3): Lower, Gorge, Middle, Tributary and Upper (Cunjak et al., 1993). Another stream is also studied, Otter brook. Five habitats type were identified in the Catamaran brook (Table. 1).

Table 1: Characteristics of the five habitat types present in Catamaran brook (Source : Cunjak et al., 1993, Girard et al., 2004).

Habitat type	Depth	Water flow	Substrates
Flat	Shallow < 46 cm	Slow < 15 cm.s ⁻¹	Small-medium particle sizes Homogeneous
Run	Deep >23 cm	Fast >15 cm.s ⁻¹	Heterogeneous substrate sizes
Bedrock	Deep >23 cm	Fast >15 cm.s ⁻¹	Bedrocks
Riffle	Shallow <23 cm	Fast >15 cm.s ⁻¹	Rubbles and Boulders Heterogeneous
Pool	Deep >46 cm	Slow < 15 cm.s ⁻¹	Sands and silts

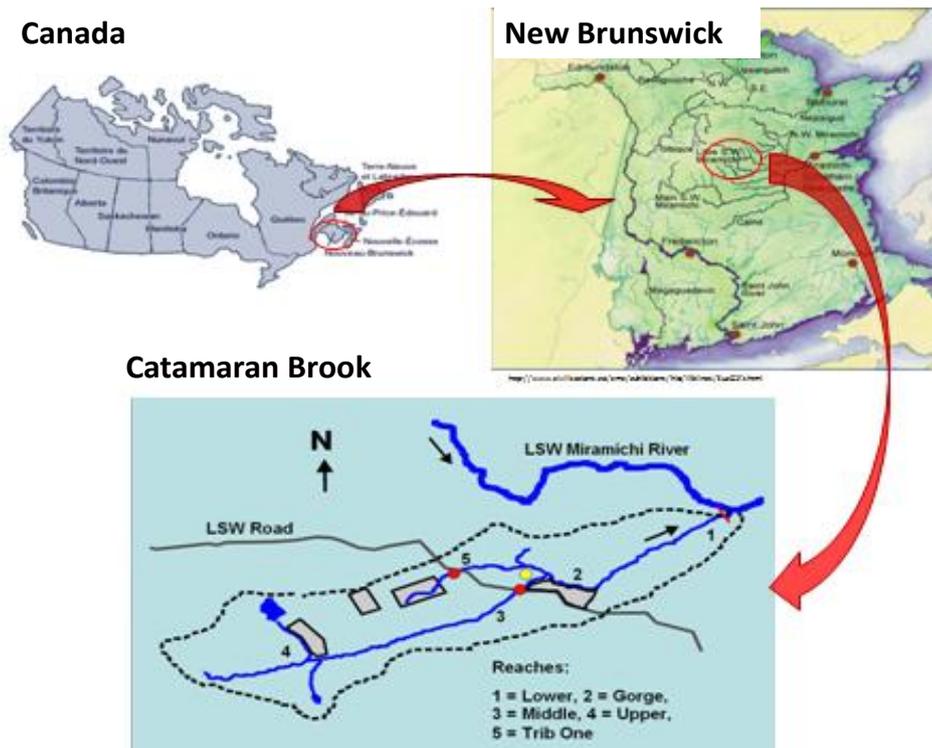


Fig 3: Situation of Catamaran brook in New Brunswick, Canada. The bottom panel gives the positions of the 5 reaches sampled in the brook.

Sampling sheme and data

Between 1990 and 2010, between 9 and 31 sites were sampled, each year by electrofishing during autumn (from late September to early December) for three age classes (0+, 1+ and 2+). Each site is a section of the stream identified by a reach in the Catamaran watershed (Lower, Gorge, Middle, Tributary, Upper or Otter Brook), a habitat type (Pool, Flat, Bedrock run, Run or Riffle). The wetted surface of each sampling site is measured (between 14 and 262 m²).

The sampled section was blocked with barrier nets and electrofished with a minimum of three successive removals and a maximum of five. After all the catches, fishes were released back into the sections from which they were captured.

The whole data set for successive removal data consists in the number of fish caught during the 3 to 5 successive removals for a total of 1479 sampling units. Each sampling unit is identified by the indices (a, h, y, r, k) :

- $a = 1, 2, 3$ stands for age classes (0+, 1+, 2+ respectively)
- $h = 1, \dots, 5$ stand for the 5 habitat types (Bedrock run, Riffle, Run, Flat and Pool, respectively)
- $y = 1, \dots, 21$ stand for years (1990-2010)
- $r = 1, \dots, 6$ stand for the 6 reaches (Lower, Gorge, Middle, Otter brook, Upper and Tributary, respectively)
- $k = 1, 2$ for the repetition per (a, h, y, r) (some year, 2 different sites with the same habitat type were sampled in the same reach).

For each sampling unit (a, h, y, r, k) , the number of fish caught during the j^{th} pass ($j = 1, \dots, 5$) is noted $C_j(a, h, y, r, k)$.

The total surface areas of all the reaches and of each habitat in each reach were also measured (Table. 2).

Table 2 : Surface Area (m²) and range of percentage of surface area sampled (during the 21 years) per reach and habitat type.

	Gorge	Lower	Middle	Upper	Total
FL	4269 0%	3662 5-12%	3747 2.5-8%	1416 0%	13094 2-5%
PO	4375 0%	3085 3-6.5%	5804 1-3%	1792 1-6%	15056 0.6-2.5%
RI	1942 4-11%	5741 2-8%	9723 0.4%	1908 0%	19314 1.5-4.5%
RU	5786 2.5-6%	18912 0.5-2%	14627 2%-6%	3397 4-8.5%	4272 0.4-3%
Total	16371 1.5-4.5%	31399 0.8-4%	33902 0.4-2.5%	8514 2-4.5%	90186 1-3%

Environmental data such as the water discharge and the spawners abundance are also available (Cunjak *com pers.*). The water discharges (m³.s⁻¹) are recorded at a hydrometric station in the Middle reach between 1989 and 2010 and a simple statistical model allows to obtain the discharge at the stream mouth (see Mitchelle and Cunjak, 2007a). The number and the age (grilse or Multiple Sea Water (MSW)) of spawners were determined from a fish counting fence located near the mouth of the brook between 1990 and 2008. Daily temperature data are also available for the period 1992-2010.

Hierarchical Bayesian Model

Available row data essentially consists in the number of fish caught during the five successive removal pass in each sampling site. These data cannot be interpreted as such for ecological inferences, but must be integrated within a statistical modelling approach to provide inferences on the density and total number of fish at different scales (reaches and the whole river stretch). A Hierarchical Bayesian Model was built to reach this goal, as sketched in Figure 4.

In the sequel, the probability distribution function f of the variable V conditionally on parameters ϕ will be noted $V \sim f(\phi)$

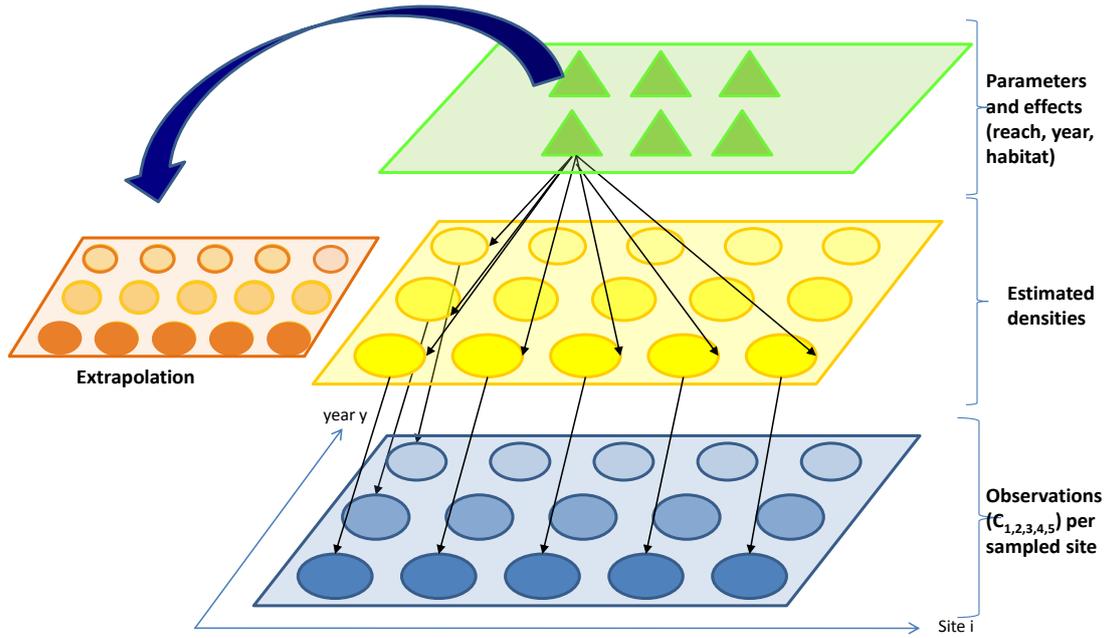


Fig 4 : Scheme of the hierarchical Bayesian Model for integrating successive removal data. The first layer (blue) represents the number of fish caught in the different sites each year. The second layer (yellow) symbolizes the estimated density for each site. The third layer corresponds to the parameters and effects estimated thanks to the data. Black arrows represent the modeling process and the statistical inference follows the reverse way.

Capture model for one site (a, h, y, r, k)

A classical model using binomial distribution was supposed for the successive removal experiment in each site (Wyatt 2002; Rivot et al. 2008; Brun et al. 2011).

Classical hypotheses were supposed to be verified: (1) During the electrofishing removal experiment, because of the small amount of time between the different passes, the population in each site is supposed to be closed in the sense that no emigration, immigration or natural mortality occur between the different passes; (2) The probability of capture may vary between sampling units, but all fish in the sampling unit have the same probability of capture and behave independently; (3) The probability of capture is considered constant between successive removal pass. Under these hypotheses, catches at each pass j , $C_j(a, h, y, r, k)$ can be modelled through a binomial distribution with initial number of fish present before the pass, denoted $N_j(a, h, y, r, k)$ and a probability of capture $p(a, h, y, r, k)$:

$$(1) \quad C_j(a, h, y, r, k) \sim \text{Binomial}(N_j(a, h, y, r, k), p(a, h, y, r, k))$$

Before the first pass, the initial number of fish in each sampling unit is denoted $N_1(a, h, y, r, k)$. The number of fish present in each site before the pass j (for $j = 2, \dots, 5$) is defined as .

$$(2) \quad N_j(a, h, y, r, k) = N_{j-1}(a, h, y, r, k) - C_{j-1}(a, h, y, r, k)$$

The initial number of fish $N_1(a, h, y, r, k)$ depends on the fish density (fish.m⁻²) $d(a, h, y, r, k)$ and the surface area of the sampling sites $S(a, h, y, r, k)$. We suppose that the fish are randomly distributed in the whole river. Then, $N_1(a, h, y, r, k)$ will follow a Poisson distribution with expected mean $\lambda(a, h, y, r, k) = d(a, h, y, r, k) \times S(a, h, y, r, k)$:

$$(3) \quad N_1(a, h, y, r, k) \sim \text{Poisson}(\lambda(a, h, y, r, k))$$

Hierarchical structure between sites

The probability of capture is known to be rather high and to vary slightly with environmental factors such as water level, discharge, fish length, However, the variability remains weak and it is hard to define systematic rules for it.

By contrast, the variability of fish density between the sampling units (a, h, y, r, k) is much higher. Population dynamics is highly variable between years (variability of spawners abundance, or survival rates ...). Spatial variability is likely to be correlated with accessibility of river section and with riverine physical habitat (Cunjak et Therrien, 1998; Bardonnnet et Baglinière, 2000).

Consistently with this background, a hierarchical structure was built to capture the variability of the probability of capture and of the density.

Probability of capture

A fully exchangeable hierarchical structure between all sampling units was built on the probability of capture (no systematic effects of year, habitat, reach, or age class are modelled a priori). For computational convenience, the hierarchical model was built in the *logit()* scale. $z(a, h, y, r, k) = \text{logit}(p(a, h, y, r, k))$ follows a Normal distribution with a mean μ_z and a standard deviation $\sigma_z = sd_z^2$:

$$(4) \quad z(a, h, y, r, k) \sim \text{Normal}(\mu_z, \sigma_z)$$

A mixture hierarchical model for the Density

A Gamma distribution for the density

The densities $d(a, h, y, r, k)$ were considered to be partially exchangeable and follow a Gamma distribution conditionally on the shape $r_d(a, h, y, r, k)$ and the inverse scale $\mu_d(a, h, y, r, k)$. These two parameters were dependent on the expected mean $E_d(a, h, y, r)$ and on the coefficient of variation $CV_d(a, y, r)$:

$$(5) \quad \begin{cases} d(a, h, y, r, k) \sim \text{Gamma}(r_d(a, h, y, r, k), \mu_d(a, h, y, r, k)) \\ r_d(a, h, y, r, k) = \frac{1}{CV_d(a, y, r)^2} \\ \mu_d(a, h, y, r, k) = \frac{1}{E_d(a, h, y, r) \times CV_d(a, y, r)^2} \end{cases}$$

A mixture model to capture the variability in accessibility

The spatial distribution of A. salmon juveniles is highly heterogeneous between the 5 different reaches (Cunjak et Therrien, 1998). For instance, Tributary and Upper reaches are never colonized by spawners. As a consequence, the 0+ density is always near 0. However, because of migrations during the juvenile phase, the density of 1+ and 2+ fish can be slightly positive for some years. In the Middle reach, colonisation by spawners is generally limited by numerous beaver dams which persist for many years after construction. The density of 0+ fish could be positive or null depending upon the accessibility conditions. But once again, because of migration of older juveniles, density of 1+ and 2+ fish are almost always positive.

As a consequence of this spatio-temporal variability related to the accessibility, the distribution of the densities between all sampling units (a,h,y,r,k) has a peak near very low densities (near 0, although 0 cannot be totally excluded because of randomness induced by binomial sampling), completed by a classical dome shaped distribution for other positive values. By analogy with zero-inflated models, a mixture model was built to capture this variability.

Variability in accessibility was captured through the Boolean variable $B(a, y, r)$ indexed by the age, year and reach. $B(a, y, r) = 1$ is associated with a non negligible density. The ecological interpretation is that the colonisation of the reach was possible. $B(a, y, r) = 0$ is associated with density near 0, because the colonization was not possible.

The values of B were not fixed in the model but estimated conditionally upon the catches data. The $B(a, y, r)$ were a priori drawn in a Bernoulli distribution with parameters $p_B(a, r)$ which are the average probability that the reach r is colonised by the age class a (informative priors were set on the $p_B(a, r)$'s ; see after and Table 3):

$$(6) \quad B(a, y, r) \sim \text{Bernoulli}(p_B(a, r))$$

Finally, the expected mean and the coefficient of variation of the Gamma distribution for the density (eq. 5) were then defined following the classical formula for a mixture model:

$$(7) \quad \begin{cases} \text{Log}(E_d(a, h, y, r)) = B_d(a, y, r) \times E_d^1(a, h, y, r) + (1 - B_d(a, y, r)) \times E_d^2(a, h, y, r) \\ CV_d(a, y, r) = B_d(a, y, r) \times CV_d^1 + (1 - B_d(a, y, r)) \times CV_d^2 \end{cases}$$

- $B(a, y, r) = 1$

The expected mean of the density was modeled in the log-scale as a linear combination of $\beta_d^1(a, y, r)$ and $\alpha_d(a, h)$

$$(8) \quad \log(E_d^1(a, h, y, r)) = \beta_d^1(a, y, r) + \alpha_d(a, h)$$

The habitat effect (interaction Age x Habitat ; Cunjak et al., 1993, Girard et al., 2004, Rivot et al., 2008) was introduced through additional fixed effects $\alpha_d(a, h)$. The variability of the density due to other factors was introduced through a random effect $\beta_d^1(a, y, r)$ standing for the interaction Age x Reach x Year. The $\beta_d^1(a, y, r)$'s were normally distributed with unknown expected mean and standard deviation depending upon the age classes, denoted $\mu_\beta(a)$ and $\sigma_\beta(a)$ respectively :

$$(9) \quad \beta_a^1(a, y, r) \mid \sim \text{Normal}(\mu_\beta(a), \sigma_\beta(a))$$

The residual variability of the density was modelled through a coefficient of variation CV_d^1 (Table. 3)

- $B(a, y, r) = 0$

This case corresponds to very low (near 0) density of juveniles in the reach r the year y . The log of the expected mean of the density was set equal to a fixed effect $\beta_a^2(a)$:

$$(10) \quad \log(E_d^2(a, h, y, r)) = \beta_a^2(a)$$

Summing up the model specification in a Directed Acyclic Graph.

A scheme of the full hierarchical model in the form of a Directed Acyclic Graph (Lunn et al. 2009) is given in Figure 5.

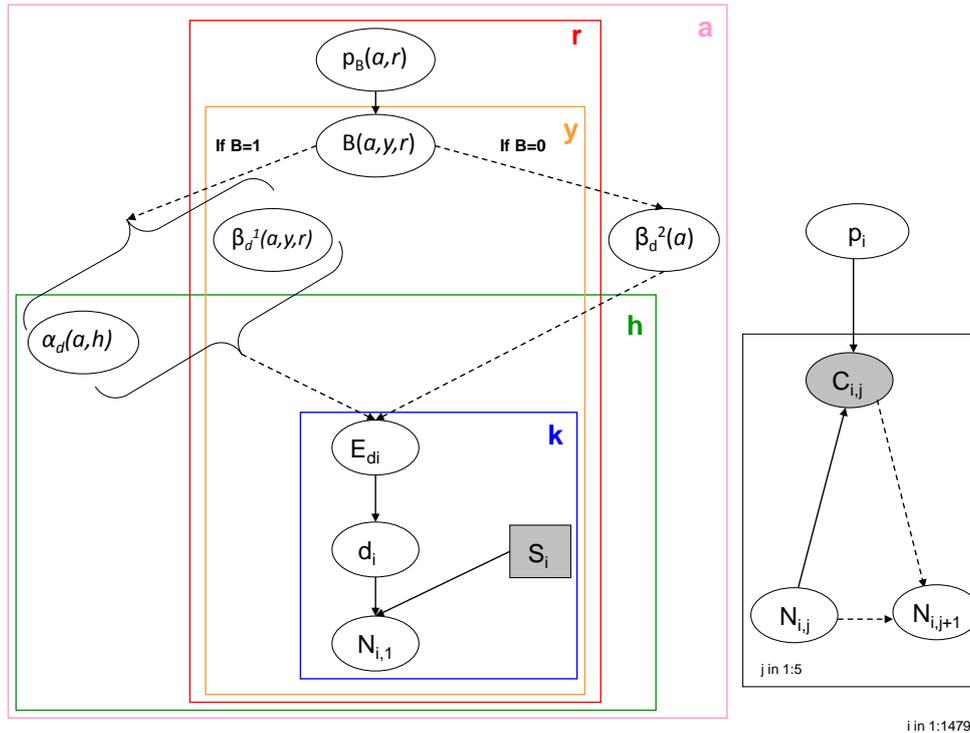


Fig 5 : Simplified Directed Acyclic graph (DAG) for the model

Prior specification for free parameters

Prior distributions were assigned to all free parameters (those that were not conditioned by any quantity of the model). For most of them, rather diffuse (weakly informative) prior distributions were used, in order to let the Bayesian posterior inferences reflect the information brought by the data (Table. 3).

More informative Beta priors distributions were set on the $p_B(a, r)$'s (Table 2). Expert knowledge about the accessibility of the reaches was introduced for some reaches. The Lower and the Gorge reaches were almost always colonised; this was introduced through a

Beta(4,1) prior distribution. The Upper and the Tributary reaches were never colonised (Cunjak *et al.*, 1993). This knowledge was introduced through a Beta(1,4) distribution. In the case of the Middle and the Otter brook reach, a high variability between year and age class due to beaver dams was observed, then a diffuse Beta(1,1) prior was used.

A constraint was introduced to avoid confusion between the habitat effects on the density: For each age class a , the sum of the $\alpha_d(a, h)$ was set to 0.

Table 3: Prior distribution of free parameters

Parameters	Distribution
μ_z	$\sim Normal(0,100)$
sd_z	$\sim Uniform(0,10)$
CV_d^1	$\sim Uniform(0,5)$
For $r = 1, 2$ $p_B(a, r)$	$\sim Beta(4,1)$
For $r = 5, 6$ $p_B(a, r)$	$\sim Beta(1,4)$
For $r = 3, 4$ $p_B(a, r)$	$\sim Beta(1,1)$
$\mu_\beta(a)$	$\sim Normal(0, 100)$
$sd_\beta(a)$	$\sim Uniform(0,10)$
$\alpha_d(a, h)$	$\sim Normal(0, 100)$ (sum to 0)
$\beta_d^2(a)$	$\sim Normal(0, 10)$

Extrapolation of the number of fish to the whole Catamaran watershed

The HBM framework was used to extrapolate fish densities and number at the scale of the reaches and finally for the whole watershed, accounting for uncertainties. Unfortunately, because the surface areas of all habitat types were not available for the Tributary and Otter Brook, the prediction could only be applied for 4 of the 6 reaches.

The extrapolation method proposed by Brun *et al.* (2011) was used. For each combination of factor *Age Class* \times *Habitat Type* \times *Year* \times *Reach*, the number of fish on the whole reach was drawn in the posterior predictive Poisson distribution:

$$(11) \quad N'(a, h, y, r) \sim Poisson(S_t(h, r) \times E_d(a, h, y, r))$$

where $S_t(h, r)$ is the total surface area for the reach r and habitat type h , (Table .2) and $E_d(a, h, y, r)$ is the expected mean of the Gamma distribution given in eq. 7 drawn in its posterior predictive distribution. The number of fish estimated from the sites sampled was not considered in this approach but the small surface area sampled compared to the total surface area (sampling rate is about 3%) minimizes the effect of such an approximation.

The total number of fish in each reach was obtained by summing the $N'(a, h, y, r)$ over all habitat types, and the number in the whole watershed (except Tributary and otter brook) was obtained by summing the $N'(a, h, y, r)$ over all habitat type and reaches.

The time series of the ratios (Number of 1+ / Number of 0+) and (Number of 2+ / Number of 1+) were computed to provide insight to the population dynamics of the juvenile

life stages. Under the hypothesis of a close population (no emigration and immigration at the scale of the brook), these ratios provide estimates of the average annual (Autumn year $t \rightarrow$ Autumn year $t+1$) survival rates of juveniles. The ratio (Number of 2+ / Number of 1+) provides an estimate of the combination of both survival and emigration as downstream migration outside the Catamaran brook occurs between 1+ and 2+ fish (Cunjak *et al.*, 1993)

Posterior Checking

Posterior predictive checking was proposed to check the model consistency. Posterior checking aims at using the a posteriori fitted model to replicate new data and to check if the replicated data looks similar to the observed ones. Here, posterior check was made at the level of the latent layer of the densities, as the modelling of the spatio-temporal distribution of the density is the most critical part of the model.

Replicated densities in each sampling unit (a, h, y, r, k) were drawn in their posterior predictive distribution :

$$(12) \quad d'(a, h, y, r, k) \sim \text{Gamma}(r_d(a, h, y, r, k), \mu_d(a, h, y, r, k))$$

with parameters $r_d(a, h, y, r, k)$ and $\mu_d(a, h, y, r, k)$ also drawn in their posterior distribution. The replicated densities were compared to the fitted densities (no observed densities are available as the densities are in a latent layer of the hierarchical structure) via the log of the ratio $(d(a, h, y, r, k) / d'(a, h, y, r, k))$.

Bayesian Computations

The joint posterior distribution of all unknown quantities (parameters, latent variables and missing data) was estimated through Markov Chain Monte Carlo (MCMC) simulations (Robert, 1996; Gelman *et al.*, 2004) thanks to the software OpenBUGS[®] V3.1.2 (Lunn *et al.* 2009) and the R-library BRugs[®] V 0.6-1. Three independent chains were used and the first 8000 iterations were discarded. Then 40 000 iterations were realized and the convergence of the MCMC chains for the model parameter was tested via the Gelman Rubin diagnostics as implemented in OpenBUGS[®]. The duration of simulation is about 14 hours.

Results

Posterior distribution of key parameters and latent variables, and checking of key modelling hypotheses

Capture probability

Posterior estimates of the overall mean of the capture probability (Table 4) shows that on average, 72.6% of the fish was captured at the first pass during electrofishing sampling. The average efficiency of electrofishing is rather high, which is consistent with previous results (Rivot et al. 2008; Cunjak et al., 1993). The posterior predictive distribution of the capture probability p_{pred} in Table 3 which represents what could be predicted from the hierarchical model for a not sampled site has a posterior mean at 0.71 with 2.5% chances to be below 0.46 and 2.5% chance to be greater than 0.89. The between site variability of the capture probability is rather low, as pointed out by the posterior estimate of $\sigma_p = sd_p^2$ which is rather low (Table. 4).

Table 4 : Main statistics of the posterior distribution of parameters for the hierarchical structure on capture probability. $\mu_p = \text{logit}(\mu_z)$

	Mean	Standard deviation	2.5%	Median	97.5%
μ_p	0.726	6.2×10^{-3}	0.714	0.726	0.738
sd_z	0.642	7×10^{-3}	0.627	0.642	0.658
p_{pred}	0.713	0.113	0.459	0.727	0.893

Figure 6 shows the estimates of probability of capture for all sampling units drawn as a function of age classes, reaches, habitat type and year. Most of the posterior estimates are between 0.70 and 0.75, which is consistent with the overall mean μ_p .

Results also highlight that a variability in the estimates of p exists, but also that no systematic effects of factors Age, Reach, Habitat type and Year can be detected. Results support the exchangeable hierarchical modelling hypothesis that was set on the p 's. They also mean that introducing systematic effects of Age, Reach, Year or Habitat type on the p 's would have provided very limited improvements to the model.

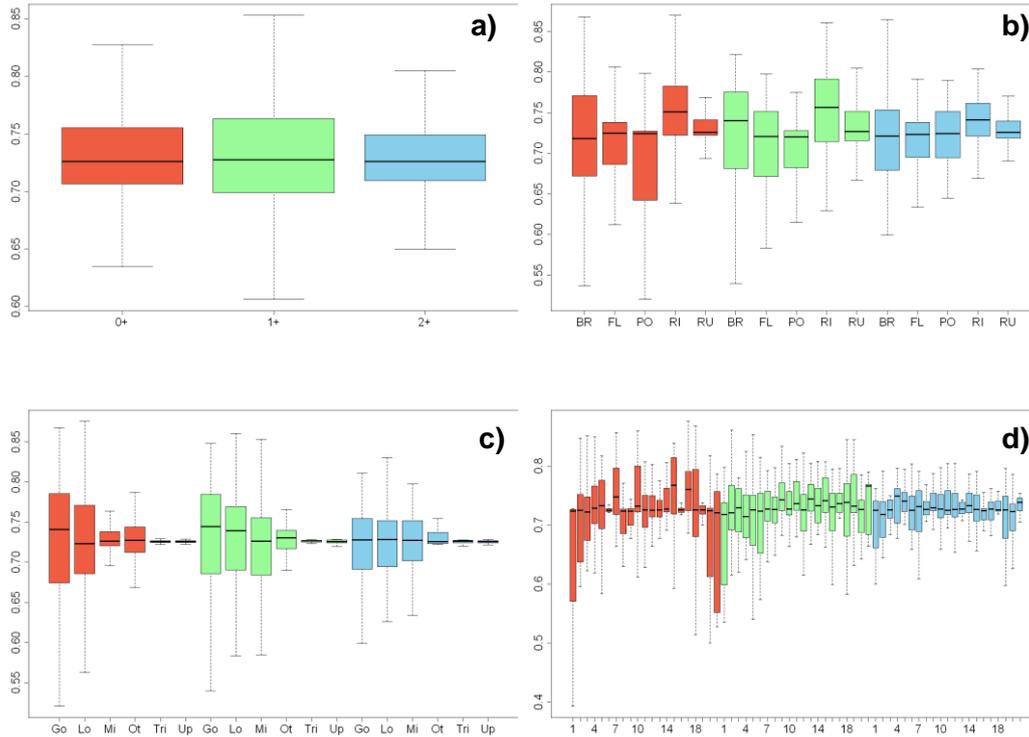


Fig 6 : Probability of capture of all sampling units drawn as a function of a) Age, b) Reach and Age, c) Habitat and Age, d) Year and Age. Boxplots represents the between units variability of point estimates (posterior median) of p . Red, green and blue represent respectively 0+,1+ and 2+

Correlation between p and d ?

During an electrofishing experiment, if we make abstraction of the Binomial sampling noise, the number of fish caught at the first pass, C_1 (i.e. the observation at hand) results from the product of the initial number of fish in the site N_1 , with the probability of capture p . Hence, in the point of view of inferences, a certain confusion may exist between the fish density and the probability of capture as an infinite number of values for the unknown (N_1, p) can lead to the same observation C_1 .

To track for such statistical identification issues, the statistical correlation between capture probability and density for all sites was assessed graphically. Results didn't reveal any spurious negative correlations between p and d , as illustrated with two examples in Fig. 7. In the case a), the two parameters are well estimated. In the case b), the probability of capture could not be precisely estimated, but the density is clearly very low (between 0 and 0.02 fish.m⁻²).

Effects on density

Mixture model

The modelling approach estimates the posterior distribution of the Bernoulli (Boolean) variables $B(a, y, r)$. The value of B can be interpreted as a indicator of the presence of fish, the value 1 indicating that the density is non negligible (the reach is colonized by the juveniles) and the value 0 indicating that the density is near 0.

Figure 8 show the time series of the posterior mean of the $B(a, y, r)$'s per age classes for the 6 reaches. As announced in the Material and Method section, the Trib. And Upper reaches are never colonized by spawners, so the density of 0+ juveniles is always near 0. Otter Brook and Middle reaches are randomly colonized, and Lower and Gorge reaches are always colonized by 0+ juveniles. Surprisingly, no sector is never colonized by 1+ and 2+ fish. This will be further discussed latter in the report.

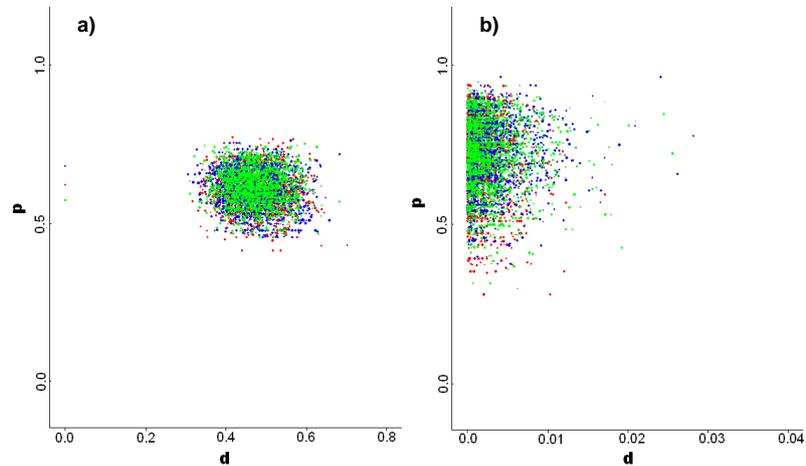


Fig 7 : Pairwise MCMC plots showing the shape of the joint posterior distribution of $p(a, h, y, r, k)$, $d(a, h, y, r, k)$ for 2 examples : a) Lower , 2006, 2+, Bedrock ; b) Middle, 2002, 2+, Run

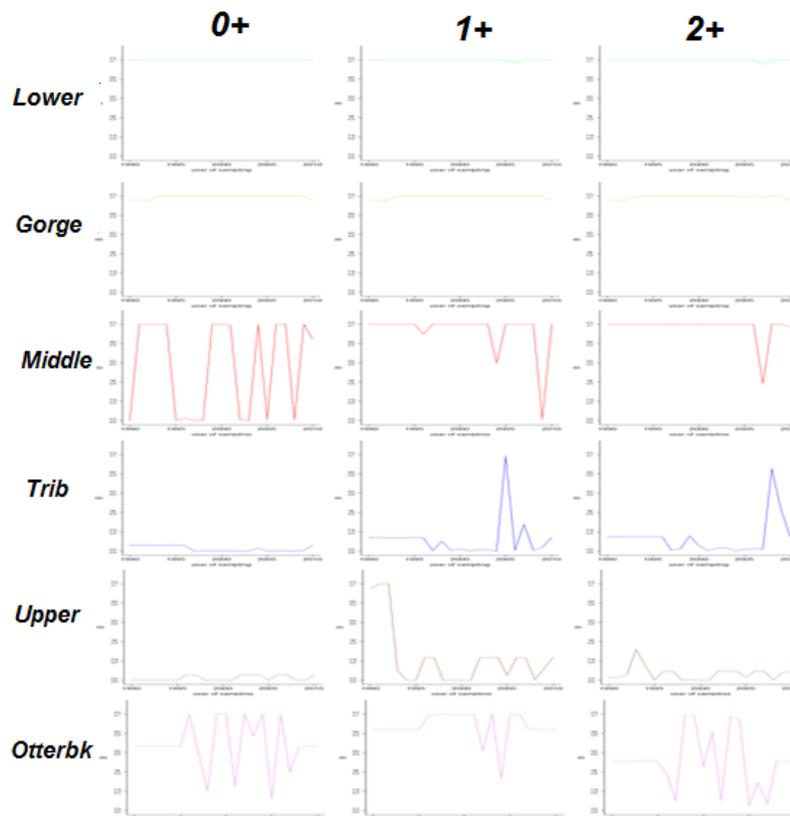


Fig 8 : Time series of the posterior mean of the Bernoulli variable $B(a, y, r)$ interpreted as an indicator of the presence of fish of the three age classes in each reach between 1990-2010. The average posterior probability that $B=1$ for the 3 ages classes are respectively $(0.97, 0.97, .0.97)$, $(0.97, 0.97, 0.97)$, $(0.56, 0.90, 0.94)$, $(0.65, 0.86, 0.51)$, $(0.04, 0.21, 0.07)$, $(0.04, 0.12, 0.13)$ respectively for Lower, Gorge, Middle, Otter Brook, Upper and Tributary.

Interaction Year, Age, Reach

Figure 9 shows the posterior distribution of the random effects $\beta_d^1(a, y, r)$ drawn as a function of Age class, Year and Reach (note that the $\beta_d^1(a, y, r)$'s were only given for the three reaches that are regularly colonized by spawners). Those parameters can be interpreted as the interaction effect of those 3 factors (on the log-scale of the density).

Results first show that average density decrease with age classes (0+, 1+ and 2+), which is consistent with the general idea of the population dynamics.

The spatio-temporal variability of the juvenile density in the watershed is quite high. No particular synchrony between reaches and between age classes appears in the time series of the $\beta_d^1(a, y, r)$'s. This result is consistent with the modelling choice that was made: an interaction term for the combination of the three factors Age class, Year and Reaches.

Again, as shown in Figure 9, for one particular age class (0+, 1+ and 2+), the overall mean of the $\beta_d^1(a, y, r)$'s (as shown by the horizontal lines in the graphs) are quite similar between the three reaches Lower, Gorge and Middle. This result is consistent with the exchangeable hypothesis across Reaches \times years that was made for the hierarchical structure on the $\beta_d^1(a, y, r)$'s.

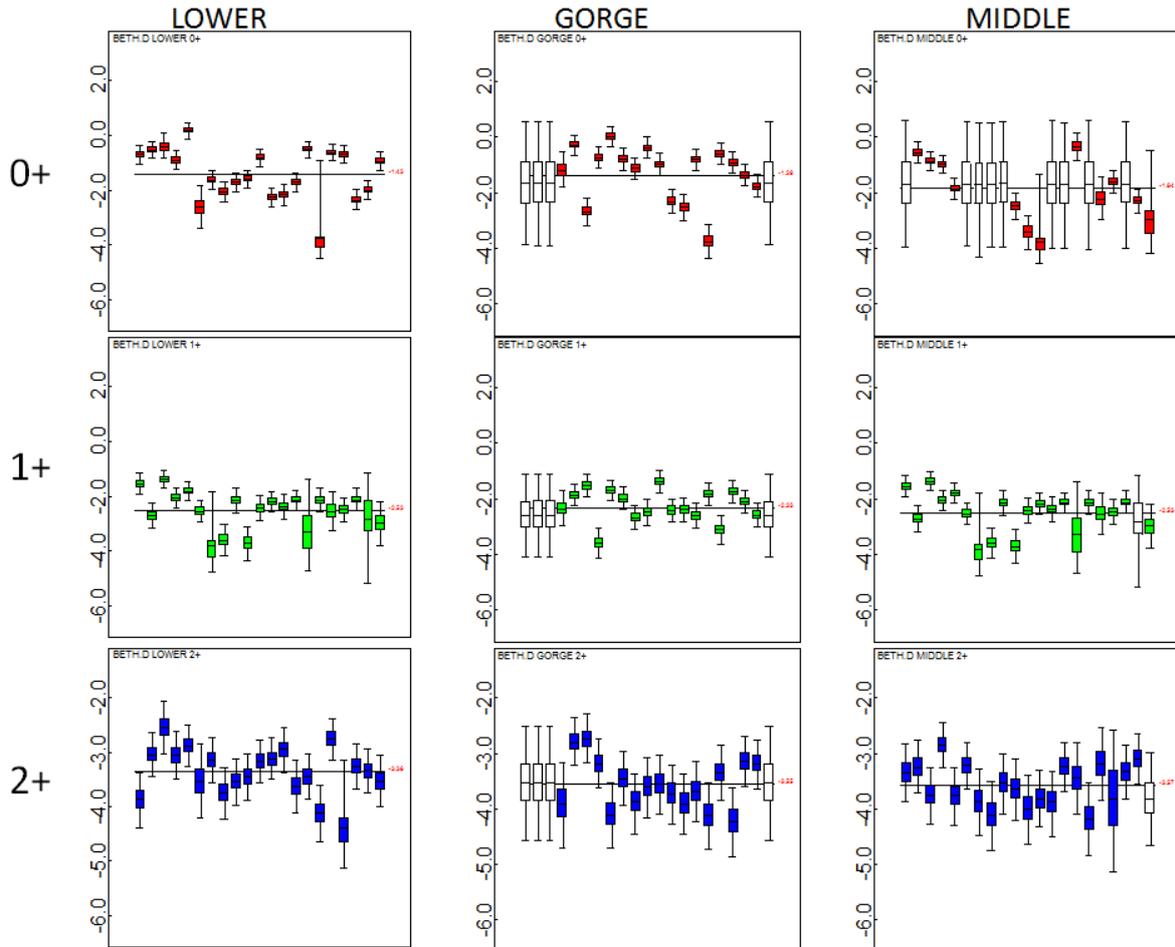


Fig 9 : Posterior distribution of $\beta_d^1(a, y, r)$ per age class and reach, the horizontal lines represent the mean of β_d^1 for each age class and reach during the 21 years (represented in the x axis). Boxplot in white font are the posterior predictive of the β_d^1 's for the combination of factors where no electrofishing data were available

Habitat preferences

The posterior distribution of the habitat effect per age class in Fig. 10 show that significant habitat effects on the density exist. These habitat effects presumably reflect age-class habitat preference (and avoidance).

Results are also in accordance with the prior assumption that the habitat effects could be specific for each age class. For instance, as shown in Fig. 10, the effect of habitat “Pool” is clearly negative for 0+ juveniles, but positive for 2+ juveniles. This might reflect a difference of habitat preference between age classes of A/ salmon juveniles.

However, these results must be balanced by the fact that the habitat and reach effects might be confounded because of the sampling scheme and unequal distribution of habitat types across all reaches. For example, all the sampling sites with bedrock run (BR) habitat type take place in the Lower and Gorge reaches where such habitats are uniquely found. These are also the two reaches where high densities of juveniles occur every year. By contrast, the samplings for Run habitats were made in every reach.

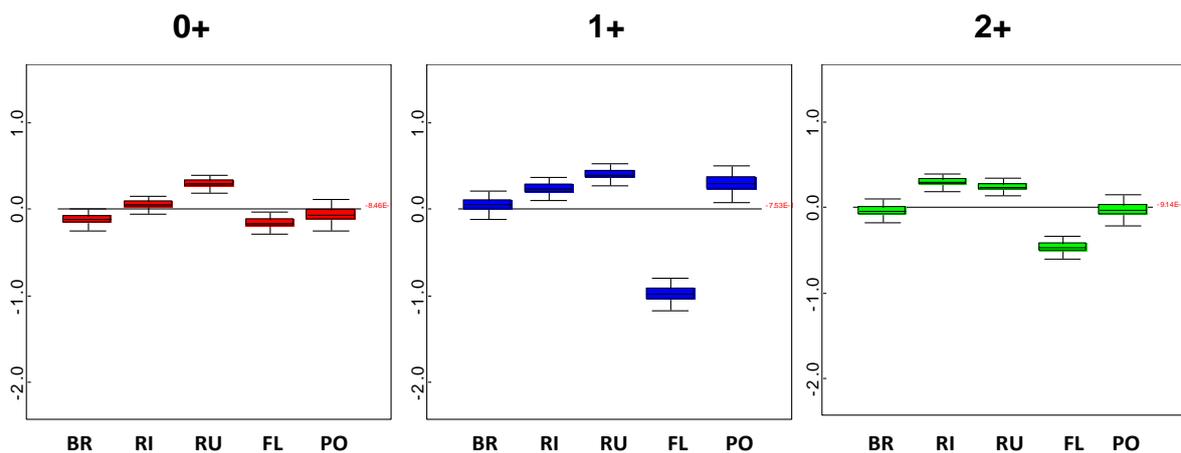


Fig 10 : Posterior distribution of habitat effects per age class $\alpha_d(a, h)$. The horizontal line reflect the above average of $\alpha_d(a, h)$ (equal to 0 by construction)

Posterior checking

The internal consistency of the model was assessed by contrasting the fitted densities with the replicated densities in each sampling site. This was done by looking at the distribution of the logarithm of the ratio $d(a, h, y, r, k) / d_{rep}(a, h, y, r, k)$. Figure 11 shows that these ratios are rather well distributed around the value 0, indicating no strong average discrepancy between the replicated densities and the fitted density, although the ratio has a huge range of variation.

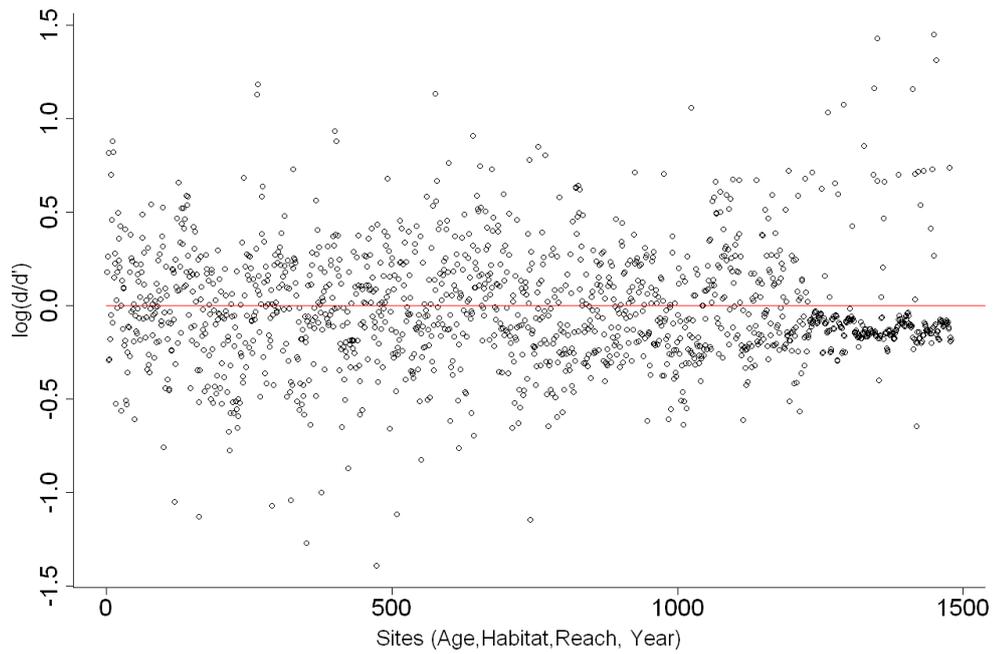


Fig 11 : Comparison between $d(a, h, y, r, k)$ and $d'(a, h, y, r, k)$ for all sampling sites. The red line represents the value 0.

However, results point out a failure of the mixture model to accurately predict the very low values of the densities observed in the Upper and Tributary reaches. Indeed, for the samples sites corresponding to these two reaches, an almost systematically value below 0 was observed (See the last sample sites in Fig. 11 and Fig. 12). For these reaches, results hence point out that the a posteriori replicated densities are almost all systematically greater than the fitted densities.

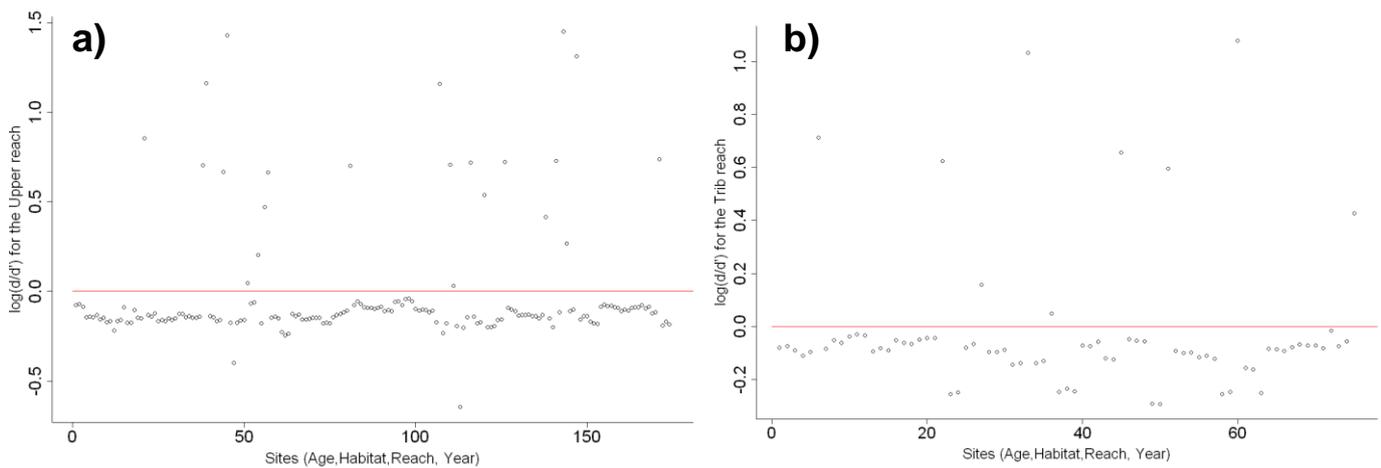


Fig 12 : Comparison between $d(a, h, y, r, k)$ and $d'(a, h, y, r, k)$ for a) the Upper reach and b) the Tributary reach. The red line represents the value 0

Abundance of juveniles and population dynamics at the scale of the Catamaran brook

Total abundance of 0+, 1+ and 2+ juveniles

A huge between year variability of the total abundance of the three age classes is observed (Figure 13). The between year variability is especially high for the 0+, and looks to be dampened for older age classes 1+ and 2+.

The estimation uncertainty is quite large. This is essentially due to the very small percentage of the wetted area sampled by electrofishing. Every year, about 3% of the entire stream surface area is sampled.

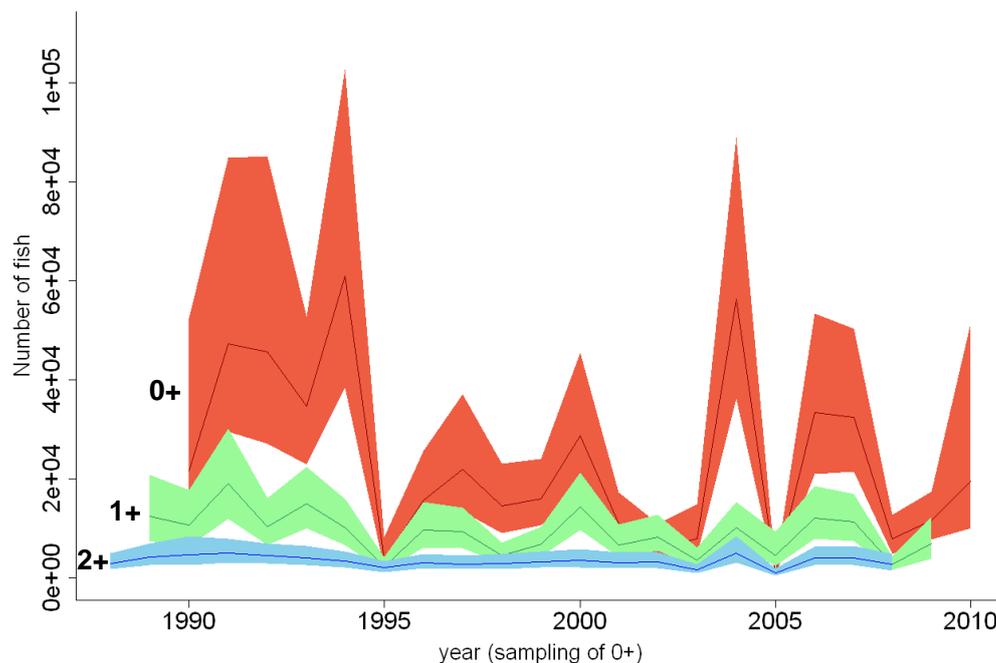


Fig 13 : Time series of estimated abundance of 0+, 1+ and 2+ salmon juveniles in Catamaran Brook, 1990-2010. 1+ and 2+ have been one (resp. two) year lagged to match with the 0+ fish, so that fish from the same cohort are aligned on the same year. For each age class, the thick line represents the posterior median and the shaded area is the posterior uncertainty (97.5% credibility interval).

First insight on the Stock-Recruitment relationship

The average density of 0+ juvenile in the whole watershed (a measure of the Recruitment) was examined in relation with the number of spawners migrating up in the Catamaran brook the year before (Figure 14a). The average density of 0+ juvenile was also examined in relation with the average water discharge during the autumn (October-November) of the spawning migration (Figure 14b).

Figure 14a pointed out that the density of 0+ seems to be related to the number of spawners. The density of 0+ juvenile seems also to depend positively on the average water

discharge in the previous autumn. But surprisingly, the number of spawners migrating up the catamaran looks independent from the average flow from October to November (Fig. 15).

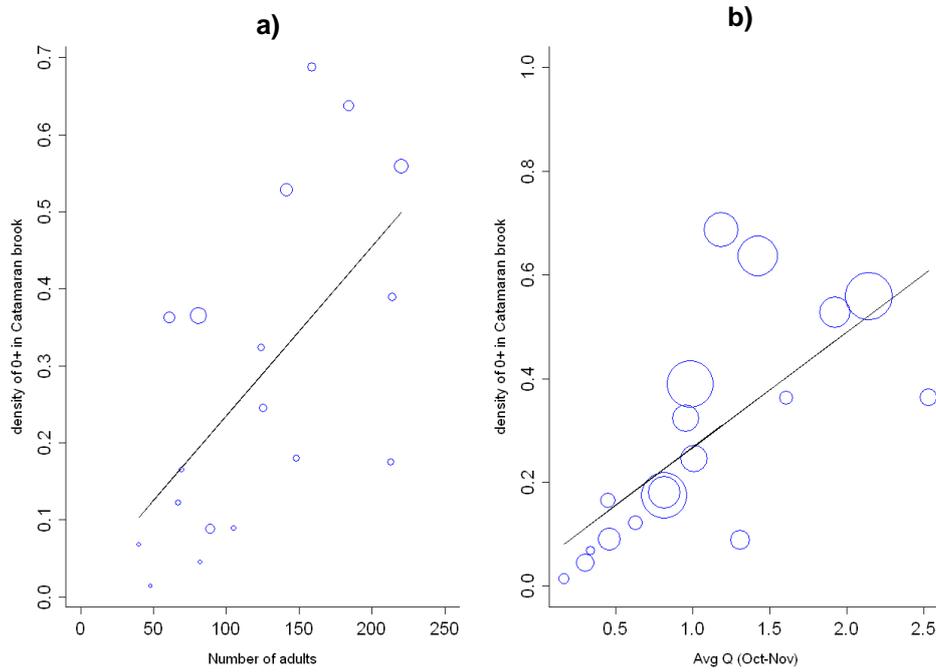


Fig 14 : Relationship between the density of 0+ and a) the number of spawners migrating up in the catamaran brook (one year lagged) (the size of the circles is proportional to the average water discharge between October and November); b) the average water flow (Oct-Nov; one year lagged) (the size of the circles is proportional to the number of spawners).

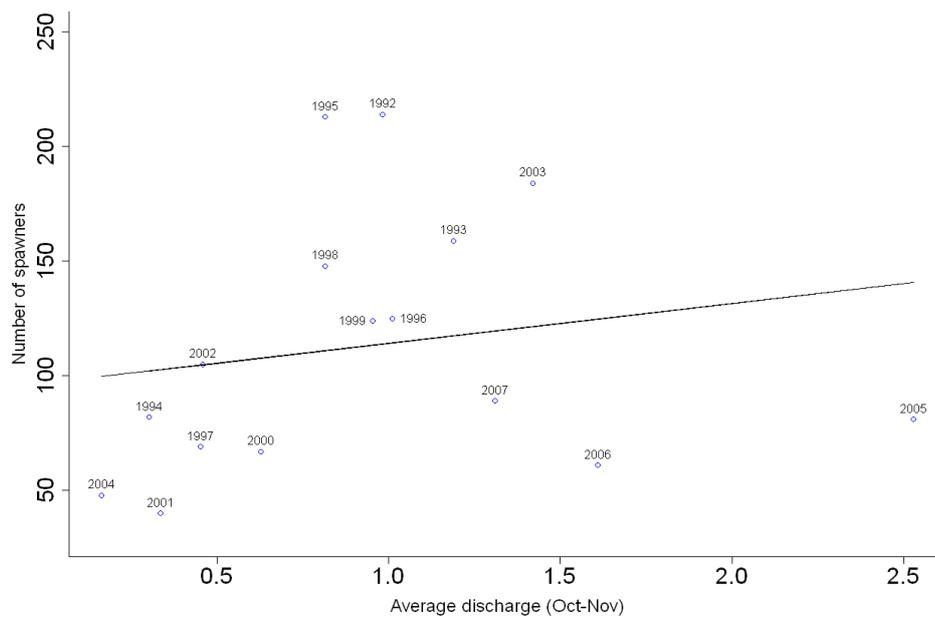


Fig 15 : Relationship between the number of spawners migrating up in the Catamarran Brook and the average water discharge (October-November).

The success of the reproduction was assessed through the analysis of the relationship between the ratio (Number of 0+/Number of spawners) and the number of spawners (Fig. 16a). Results point out that no particular pattern, except a huge random variability, is obvious in this relationship. No density dependence effect appears in the stock-recruitment relationship.

By contrast, a clear positive correlation exists between the ratio (Number of 0+/Number of spawners) and the average water discharge in autumn (Fig. 16b). Hence, the average discharge during the autumn seems to be more consequential for the recruitment than the number of spawners, suggesting a strong environmental control of the recruitment success by the average water discharge during the autumn of the spawning migration.

The recruitment and the success of recruitment were also studied in relation with other environmental variables like temperature and water discharge considered during other periods but none relationship was found.

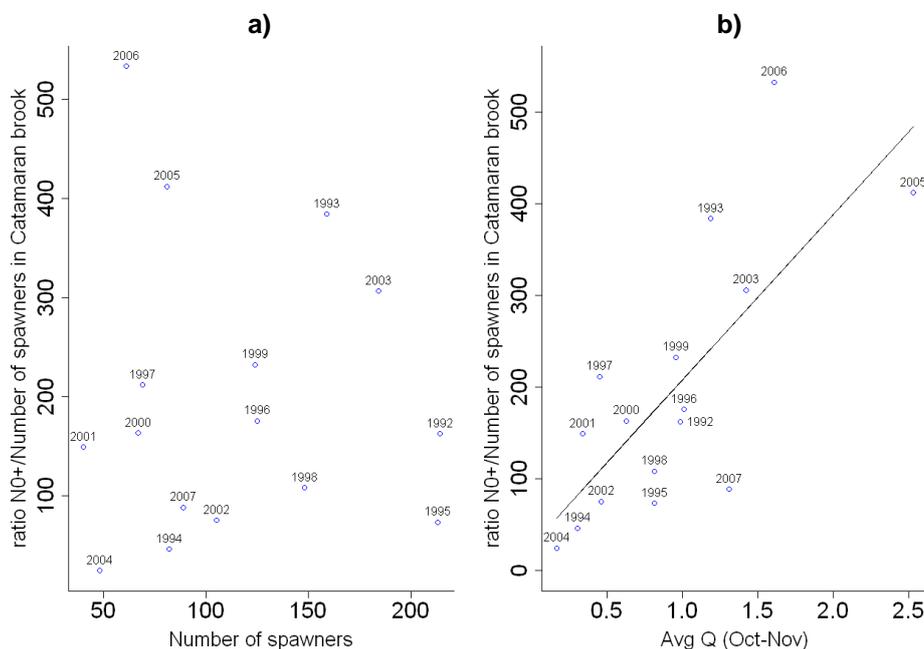


Fig 16: Relationship between the ratio (Number of 0+/Number of spawners) and a) the number of spawners ; b) the average water discharge (October-November).

Ratio N1+/N0+ and ratio N2+/N1+

The ratio (Number of 1+ / Number of 0+) (in log-scale) is below 0 (<1 in natural scale) for almost all years in the time series, except for years 2002 and 2005 for which positive values (>1 in natural scale) cannot be excluded. Clearly, estimates for years 2002 and 2005 are totally inconsistent in terms of population dynamics and may be an artefact of the sampling scheme. After excluding these two years, the average ratio is around 0.36, which is consistent with the value of an annual survival rate (Fig.17a).

The ratio (Number of 2+ / Number of 1+) (in log-scale) is below 0 (<1 in natural scale) for almost all years in the time series, except for years 1995 and 2008 for which positive values (>1 in natural scale) cannot be excluded. The average ratio calculated over the time series is about 0.44 (Fig. 17b).

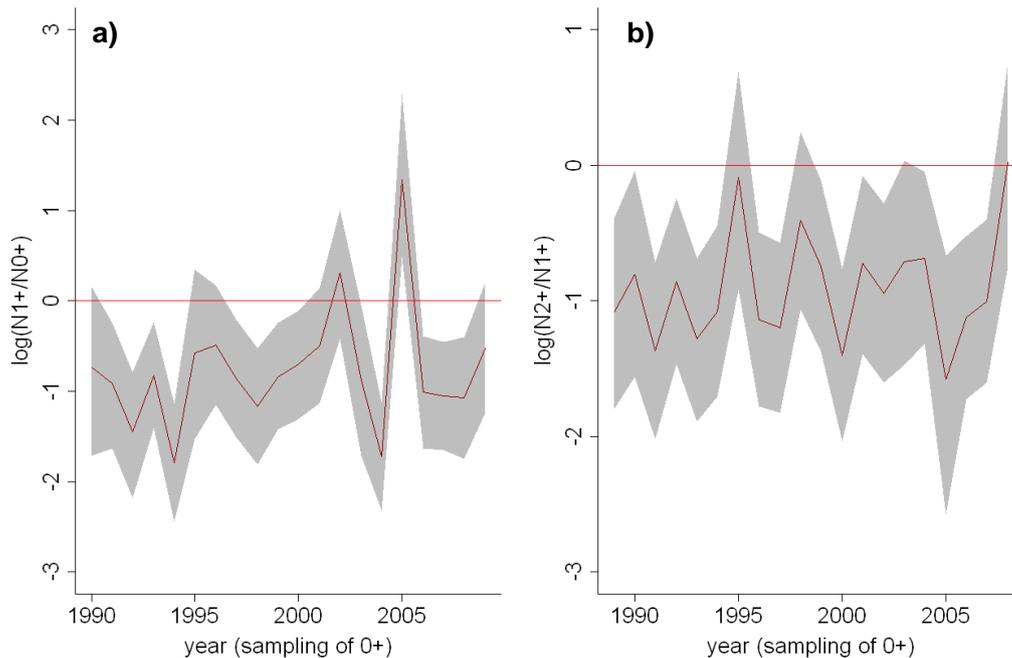


Fig 17 : Time series of posterior estimates of a) the ratio $\log(N_{1+}/N_{0+})$. 1+ fish have been one year lagged to match with the 0+ fish b) the ratio $\log(N_{2+}/N_{1+})$. 2+ fish have been one year lagged to match with the 1+ fish.. Thick line: posterior median; Shaded area: 97.5% Bayesian posterior credibility interval.

Functional role of the different reaches with regards to the population dynamics

Relative contribution of the reaches to the total abundance

The contribution of each reach to the total abundance of juvenile salmon in Catamaran Brook varies highly between years (for all age classes) (Fig.18). The patterns are also clearly different between age classes.

The greater temporal variability in the relative contribution is observed for the 0+ juveniles age classes. The Lower and Gorge reaches have the greatest contribution (contribution between 55% and 100%). These reaches are closest to the mouth of Catamaran Brook and are typically accessible to spawning adults each year. Contribution of the Middle reach is highly variable, (contribution between 0% and 50%) between years, in relation to accessibility of this sector to the spawners during the upstream migration.

Interestingly the contribution of the different reaches seem to be more and more homogeneous (between years) for 1+ and 2+ fish. Roughly speaking, the distribution of 2+ every year is about 40% in the Middle reach, 10% in the Gorge reach and 50% in the Lower reach whatever the distribution of the 0+ of the same cohort (Fig. 18, right panel). Hence, even if the Middle reach is not always colonized by 0+ fish, the functional role of this sector

of the Catamaran brook with regard to the total production of 2+ parrs is almost always very important.

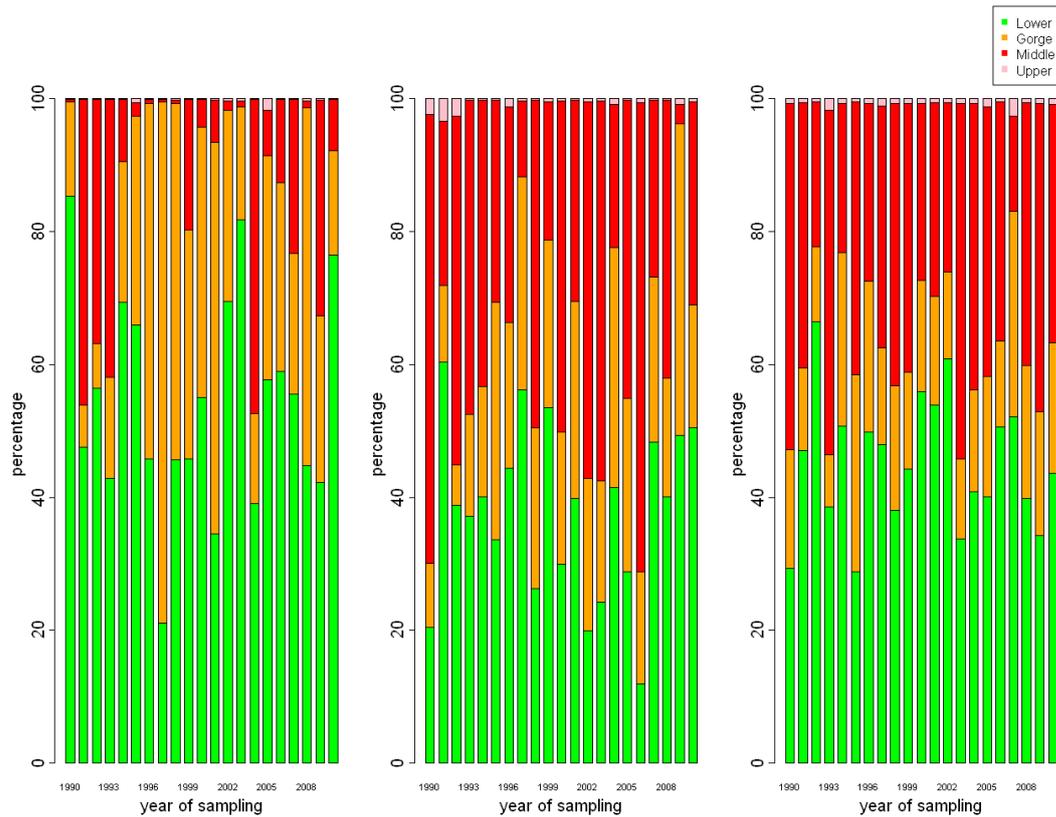


Fig 18 : Time series of the relative contribution of the 4 main reaches to the overall abundance for the three age-classes (0+, 1+ and 2+) of juvenile salmon in the catamaran brook.

Figure 19 brings another interpretation of the same results. It highlights how the contribution of the different reaches changes for 0+/1+/2+ fish of the same cohort, for 4 particular cohorts chosen to illustrate the different patterns existing in the distribution. The four examples point out differences in the spatial dynamics of the cohorts. For cohorts born in 1996 and 1997, the contribution of the Middle Reach was very low for 0+ but get higher for 1+ and 2+. This might suggest an important migration of 1+ and 2+ into the Middle Reach. By contrast, for cohorts born in 1993 and 2004, no change in the relative contribution of the 4 reaches appears when fish get older and older.

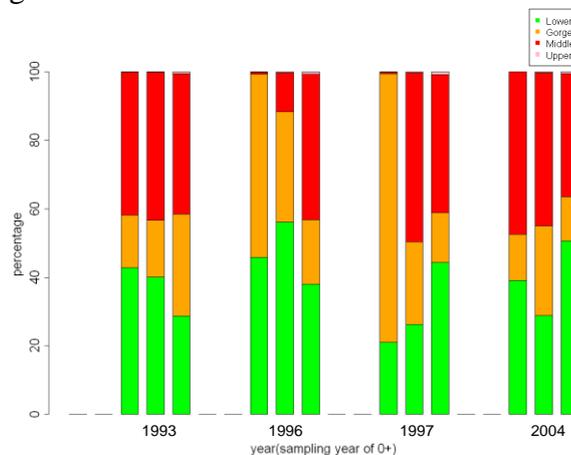


Fig 19 : relative contribution of the 4 main reaches to the overall abundance for the three age-classes (0+, 1+ and 2+) of juvenile salmon as demonstrated for four cohorts 1993, 1996, 1997 and 2004 (year of birth of 0+ juveniles).

Relationship between accessibility and environmental covariates

Logically, the huge between year variability of the contribution of the reaches to the recruitment (measured in term of 0+ juvenile abundance) highlighted in Fig. 18 is related to the variability of the accessibility to spawners (Fig. 20). In particular, for the Middle reach, years with low contribution of the Middle reach to the 0+ recruitment are in general years for which $B=0$.

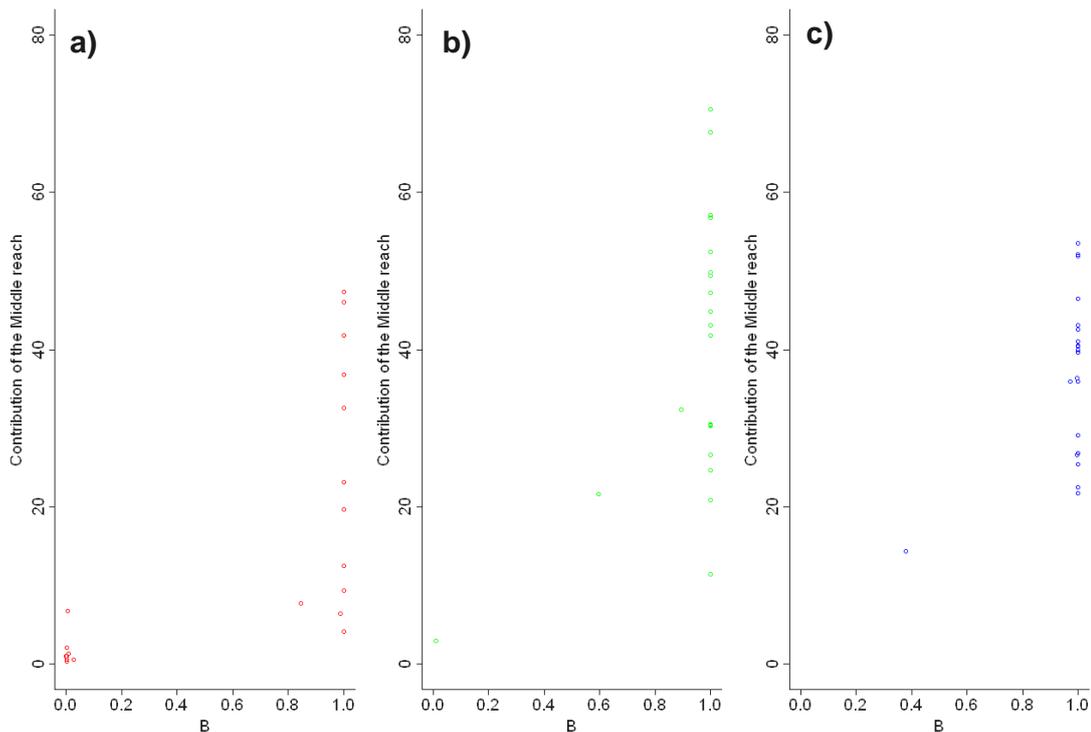


Fig 20 : Relation between the relation contribution of the Middle reach and the accessibility (B) of this sector a) 0+, b) 1+, c) 2+

The relationship between the accessibility and the water discharge can be examined further through the relationship between the indicator variable B and the average water discharge between October and November. Fig. 21b suggests that a certain threshold of water discharge seems to be necessary to allow the migration of spawners in the Middle reach. By contrast, Gorge and Lower reaches are always accessible even in low water discharge conditions (Fig. 21a).

Surprisingly, even if the recruitment success appears as positively correlated with the average water discharge in the autumn of the spawning migration (as showing Fig. 16b), no clear relationship could be found between the whole recruitment success at the scale of the Catamaran and the relative contribution of sectors which accessibility highly depends upon the water discharge conditions

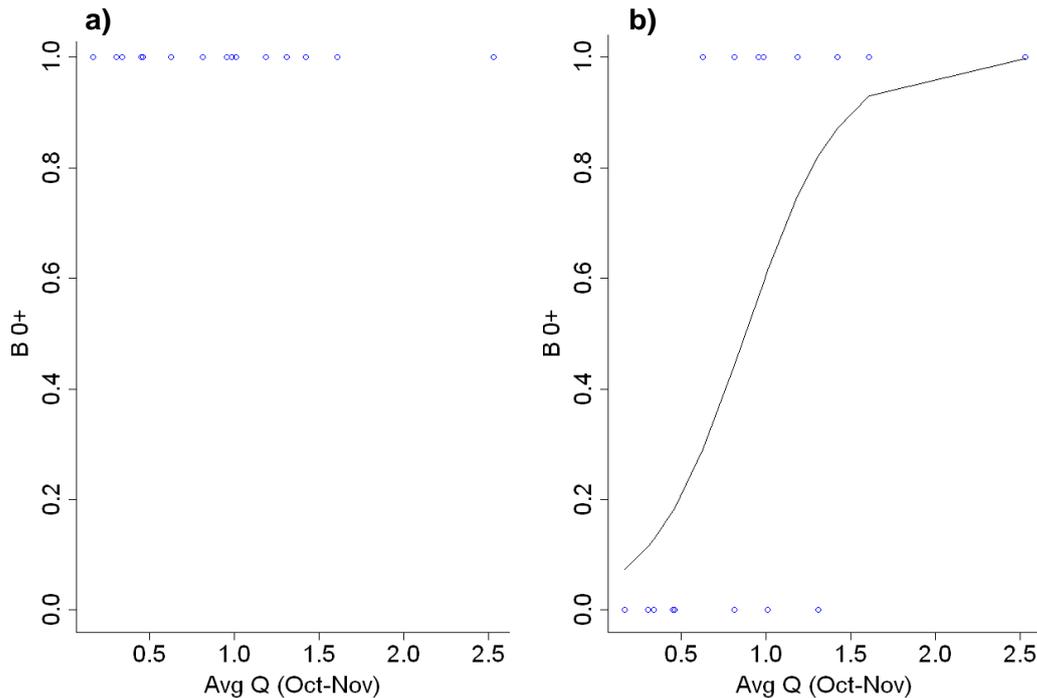


Fig 21 : Posterior median of the indicator variable B (interpreted as an indicator of the accessibility) for 0+ juveniles as a function of the average discharge (Oct-Nov) of a) Lower reach, b) Middle reach.

Evidence of an intense migration between reaches

Results also suggest that an intense migration of salmon juveniles occurs within the watershed.

For instance, the ratio (Number of 1+ fish (one year lagged)/ Number of 0+ fish) in the Middle reach is very often much greater than 1 (>0 in log scale) (Fig. 22a). Hence, for some cohorts, the density of 1+ fish in this sector is high even if the density of 0+ fish was near zero the year before. This reinforces the hypothesis of upstream migration. By contrast, the ratio N_{2+}/N_{1+} is rarely greater than 1 (>0 in log scale) (Fig. 22b), thus suggesting that the essential of the upstream migration seems to be made by the age class 1+.

The upstream migration of juvenile from the Gorge reach to the Middle reach was studied in relation with variables like temperature, water discharge during the migration period and density dependence phenomena but none relationship was found.

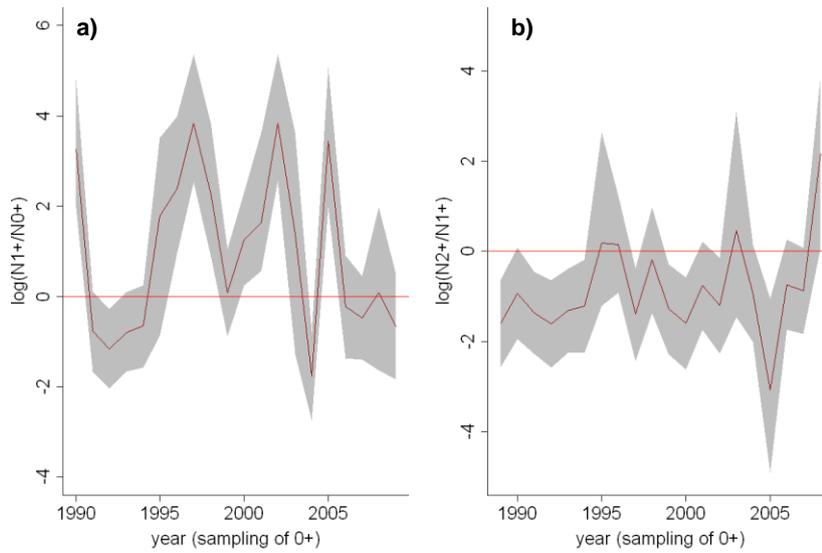


Fig 22 : Time series of the ratio a) (Number of 1+ fish / Number of 0+ fish) (in log-scale); b) (Number of 2+ fish / Number of 1+ fish) in the Middle reach. The 1+ have been one year lagged with respect to 0+ fish, and the 2+ fish have been one year lagged with respect to the 1+ fish.

Discussion

Model

Following the approaches of Wyatt (2002), Wyatt *et al.* (2003), Rivot *et al.* (2008), Dauphin *et al.* (2009) and Brun *et al.* (2011), a Hierarchical Bayesian Model was built to estimate the spatio-temporal distribution of wild A. salmon juveniles in the Catamaran brook from a large data set of successive removal data via electrofishing.

Some simplifying assumptions were made. In modelling, the capture process via electrofishing, the simplifying assumption was made that the probability of capture was constant across successive pass. This hypothesis has been criticized because a decrease in the probability of capture across the successive removal pass is likely to occur. This could arise from: (1) An heterogeneous catchability among individuals leading to the most easily captured individuals being caught first (discussed in Mäntyniemi *et al.*, 2005); (2) A change in the fish behaviour or the reduced susceptibility to electricity for fish already having experienced an electric shock (Bohlin *et al.*, 1990). In the data set analyzed in the present study, 3 to 5 removal pass are available on each sampling site. Such a data set would have allowed us to test for the decrease in the probability of capture across the successive removals, following the approach of Wyatt (2002), Dauphin *et al.* (2009) or Brun *et al.* (2011). However, this would have been in the price of increasing model complexity, but with likely only small improvement in the accuracy of population estimates.

The choice of a Poisson distribution for modeling the abundance and the capture processes is also discussable. Indeed, one of the underlying hypothesis for Poisson-Binomial model is that fish are totally randomly distributed and behave independently with regards to the capture process. Schooling or any other behavior or phenomenon leading to an aggregation of fish can lead to strong departure from this hypothesis (Linden and Mäntyniemi 2011). Using a Negative Binomial instead of a Binomial distribution for the capture process can help capturing a part of the overdispersion induced by the departure from the Poisson hypothesis (Linden and Mäntyniemi, 2011).

However, because the originality of our approach was rather on the modeling of the spatio-temporal distribution of the density (*e.g.* through the mixture approach), we rather chose to keep the Poisson-Binomial successive removal capture model simpler.

Beyond these limits, the approach offers a substantial contribution to the modelling of the abundance of salmonid juveniles in the wild from successive removal data sets.

The present work is another application illustrating how the HBM framework is powerful and flexible for processing multiyear and multisite successive removal data. It successfully enabled the joint treatment of several years of data and integration of a habitat-categorical covariate to explain spatial variability of the density of salmon juveniles. It takes advantage of the “borrowing of strength” (Rivot and Prévost 2002; McAllister *et al.*, 2004; Robert *et al.*, 2010) between sampling units to improve the estimation of the total salmon population size in the three age classes every year. The HBM framework successfully

accommodated the large data set, which contain poorly informative or even missing data for some units, in particular for the 2+ age class for which catch data are often very low. The conditional probability structure enables the transfer of information from data-rich to data-poor units, while explicitly accounting for the cross-units variability of the density and probability of capture. The posterior predictive distribution of the probability of capture conveys all the posterior information that can be used to infer the density at any site where only very low numbers of fish were caught. The model provides estimates of the population size on a whole river stretch, while accounting for the two sources of uncertainty: (i) the sampling variance due to the removal method in sampling sites and (ii) the uncertainty arising from the across-site variability of the density.

The most original part of our modelling approach is the mixture modelling approach introduced to model the spatio-temporal distribution of the density that enables to capture both very high and very low densities. Our approach has lot of similarities with zero-inflated models developed for counting data in ecology (Martin *et al.*, 2005; Sileshi *et al.*, 2009; Ancelet *et al.*, 2010; Calama *et al.*, 2011; El-Shaarawi *et al.*, 2011; Vaudor *et al.*, 2011). However, to our best knowledge, our model is the first one that merged a mixture model for the density within an integrated statistical model for successive removal data.

Moreover, the mixture model introduced a latent variable (the Bernoulli variable B in the model) with an ecological interpretation linked with the process of accessibility to certain sectors of the Catamaran brook. This latent variable takes the value 0 when the sector was not colonized by salmon juveniles and equals 1 if the sector is colonized. But the status of this variable is stochastic and the probability to be 0 or 1 is interpreted as the probability to be actually colonized. For some reaches like the Middle reach, the status of B for 0+ juveniles was successfully related to environmental factors like the average water discharge during the autumn during the migration of spawners.

Population dynamics

The HBM provides 21-years time series of abundance estimates for the three age classes of *A. salmon* juvenile in the catamaran brook. The complex hierarchical structure enabled us to assess the distribution of abundance at different spatial scales, namely the whole Catamaran watershed, the reaches within the watershed and the habitat types within reaches. This provided substantial insights to the population dynamics of the juvenile phase of the *A. salmon* life cycle in this watershed.

At the scale of the whole watershed, the time series of abundance estimates first highlighted a huge between year variability of the 0+ juveniles. This is consistent with previous knowledge on *A. salmon* population dynamics, as the recruitment of 0+ juveniles is known to be highly sensitive to environmental fluctuations (Jonsson *et al.*, 1998; Elliott 2001; Milner *et al.* 2003). Such a result is also consistent with estimates found in other watersheds. For instance, in the Oir river (Lower Normandy, France), Baglinière *et al.* (2005) and Rivot *et al.* (2008) found that the 0+ salmon densities fluctuated widely between years.

The density (fish.m⁻²) of 0+ juveniles found in the Catamaran brook was more variable and with a maximum rather high by comparison with estimates obtained for other watersheds. For instance, densities on riffle habitat types were found to be between 0.009 and 0.59 in the

Catamaran (this study) and between 0.018 and 0.38 in the Oir river (Rivot *et al.*, 2008). This difference in densities could be explained by the importance of the Miramichi River which is one of the most productive Salmon River in North America. Also, this must be balanced by the fact that although they have the same denomination, habitat types are not necessarily equivalent between such a Canadian and French river.

Habitat preferences were also found to be consistent with other studies (Baglinière and Champigneulle, 1986; Bardonnet and Baglinière, 2000; Brun *et al.*, 2011) An other interesting result is that habitat preferences revealed age-specific especially for Pool habitat which seems to be avoid by the 0+ and 1+ age classes like Flat habitat but not by 2+.

The time series of 0+ abundance estimates allowed for a first examination of the Stock-Recruitment relationship. No strong Stock (measured as the abundance of spawners) – Recruitment (measured as the abundance of 0+) was found, but an environmental control of the Recruitment success by the water discharge was pointed out. Indeed, the abundance of 0+ revealed relatively independent from the abundance of spawners the previous year. Similarly, the ratio (abundance 0+ / abundance spawners) revealed independent from the abundance of spawners. Thus, no evidence for density dependence could be found. But interestingly, the ratio (abundance 0+ / abundance spawners) revealed positively and significantly correlated with the average water discharge during the autumn of the spawning migration. No relationship was found with the water discharge in other periods of time neither with temperature.

The spatialized structure of the model also gave access to estimates of the relative contribution of each sector to the juvenile production in the whole watershed. The contribution of each reach to the total abundance of salmon juveniles in Catamaran Brook varied highly between years (Fig.18). The patterns were also clearly different between age classes.

The greatest temporal variability of the contribution between reaches was observed for the 0+ age group. This is largely related to the between year variability of accessibility during the spawning migration. Beaver dams are important factor (in New Brunswick, not in France !) that control accessibility in interaction with water discharge. Beavers will commonly build a series of accessory dams which create series of barrier to the migration (Collen et Gibson, 2001). Accessibility will be dependent upon discharge, because only high water flow can facilitate passages through the dam for salmon adapted at swimming in high flows and jumping over barrier (Collen et Gibson, 2001). In the Catamaran brook, the variability in the accessibility in relation with water discharge is particularly visible for the Middle reach, for which the indicator variable B was found positively related to the water discharge during the upstream migration of spawners.

When accessible to spawners, the Middle reach revealed productive with regards to the contribution of the total 0+ recruitment. In parallel, the ratio (0+ abundance / Spawners abundance) revealed positively correlated to the average water discharge. But surprisingly (and unfortunately), no clear relationship was found between the relative contribution of the Middle reach to the whole 0+ recruitment and the ratio (0+ abundance / Spawners abundance) indicator of the recruitment success. In other terms, although a clear control of the recruitment success (measured as the ratio (0+ abundance / Spawners abundance) by the water discharge was found, no clear functional relation could be found between the overall 0+ recruitment at

the scale of the Catamaran brook and the contribution of sectors which production revealed regulated by water discharge.

An other interesting result is that the relative contribution of the different reaches with regard to juveniles production is much less variable for older age classes (2+) than for the 0+ age class. In particular, the contribution of the Middle reach with regard to the total production of 2+ parr is always significant (40% on average), even if the reach was practically not colonized by 0+ fish of the same cohort. Three hypotheses (not exclusive) can be formulated to explain this observation: (1) A difference between the reaches in the survival rate might exist, with much higher survival rates of 0+ and 1+ fish in the Middle reach than in other sectors such as Lower and Gorge; (2) A downstream migration between the Gorge and the Lower reaches and the little southwest Miramichi for 1+ occurs, whereas no migration outside the Middle reach occurs; (3) An upstream migration of juveniles from the Gorge to the Middle reach between the 0+ and the 2+ stages.

This last hypothesis was reinforced by other results, showing that the migration between the 0+ and 1+ stages is more intense than between the 1+ and 2+ stages. Such a migration takes place in early summer (June-July) and sometimes in October (Cunjak *et al.*, 1990). Several hypotheses (non exclusive) were made to try to interpret these migrations: i) A density dependent migration (Steingrímsson *et al.*, 2003) ; ii) The effect of extreme temperatures requiring juveniles to move to find better living conditions; iii) The effect of flow during the migration period that encourages juveniles to move to reach an area richer in preferred habitat. In this study these assumptions were tested but none of them were verified. This migration has been observed by Steingrímsson *et al.* (2003) and Breau *et al.* (2007) but at the scale of individual. The impact of environmental variables on the growth rate of juvenile has already been studied (Jensen and Johnsen, 1999; Imre *et al.*, 2005) but not the impact on upstream migration of juvenile. Further analyses are needed to understand the factors controlling juvenile migration within and outside the watershed.

Prospects

The population model provides a framework for structuring further research and data collection in the near future.

Further analyses of the Stock-recruitment relationships could also be developed. Indeed, the success of recruitment has been shown as being dependent on the water flow during the period of spawners migration. A stock-recruitment model with a productivity parameters depending on environmental covariates such as the water discharge could be developed.

The model jointly analyses the three juvenile age classes 0+, 1+ and 2+. However, these age classes were considered to be independent, in the sense that the population dynamics processes (*e.g.* survival, migration) between these three juvenile age classes are not represented. Following ideas developed in Buckland *et al.* (2004), Rivot *et al.* (2004) or Buckland *et al.* (2007), one of the most promising perspectives opened by this work would be to embed the HBM model built for electrofishing data within a population dynamics model explicitly representing the population dynamics between the different age classes.

Representing the population dynamics between 0+, 1+ and 2+ in the Catamaran watershed would certainly require additional observations and data as it requires to jointly model survival and migrations within the watershed between reaches and outside the watershed. A tentative model for modeling the migrations between the Gorge and the Middle reach was proposed and fitted to the time series of abundance data. However, available information provided by series of abundance only revealed insufficient to jointly estimate survival and migration parameters.

References

- Acou, A., Rivot, E., Van Gils, J.A., Legault, A., Ysnel, F., Feunteun, E. (2011). Habitat carrying capacity is reached for the European eel in a small coastal catchment: evidence and implications for managing eel stocks. *Freshwater Biology*, 56, 5, 952-968.
- Ancelet S., Etienne M.P., Benoît H., Parent E. (2010). Modelling spatial zero-inflated continuous data with an exponentially compound Poisson process. *Environmental and Ecological statistics*, 17, 3, 347-376
- Baglinière J.L., Cahmpigeulle, A. (1986). Population estimates of juvenile Atlantic salmon (*Salmo salar*) as indices of smolt production in the R. Scorff, Brittany. *J. Fish. Bio.* 29, 467-482
- Bardonnnet. A, Baglinière, J.L. (2000). Freshwater habitat of Atlantic salmon (*Salmo salar*). *Can. J. Fish. Aquat. Sci.*, 57, 2, 497-506.
- Bohlin, T., Heggberget, T.G., Strange, C., (1990). **Electric fishing for sampling and stock assessment.** In: Cowx, I.G., Lamarque, P. (Eds.), *Fishing with Electricity. Applications in Freshwater Fisheries Management.* Fishing New Books, Oxford, pp. 112–139.
- Breau C., Weir L.K., Grant J.W.A. (2007). Individual variability in activity patterns of juvenile Atlantic salmon (*Salmo salar*) in Catamaran Brook, New Brunswick. *Can. J. Fish. Aquat. Sci.*, 64, 3, 486-494.
- Brun M, Abraham C., Jarry M, Dumas J, Lange F, Prévost E. (2011). Estimating an homogeneous series of a population abundance indicator despite changes in data collection procedure: A hierarchical Bayesian modelling approach, *Ecological Modelling*, 222, 5, 1069-1079
- Buckland, S. T, Newman, K. B., Thomas, L., & Koesters, N. B. (2004). State-Space Models for the Dynamics of Wild Animal Populations. *Ecol. Model.*, 171, (1-2), 157-175.
- Buckland, Stephan T., Newmann, K. B., Fernández, C., Thomas, Len, & Harwood, J. (2007). Embeding Population Dynamics Models in Inference. *Statistical Science*, 22, 1, 44-58.
- Calama, R., Mutke, S., Tomé, J., Gordo, J., Montero, G., & Tomé, M. (2011). Modelling spatial and temporal variability in a zero-inflated variable: The case of stone pine (*Pinus pinea* L.) cone production. *Ecological Modelling*, 222, 3, 606-618.
- Collen P., Gibson R.J. (2001). The general ecology of beavers (*Castor spp.*), as related to their influence on stream ecosystems and riparian habitats, and the subsequent effects on fish—a review. *Reviews in fish biology and fisheries*, 10, 4, 439-461
- Cunjak, R.A., Caissie, D. & El-Jabi, N. (1990) The Catamaran Brook Habitat Research Project: Description and General Design of Study. *Can. Tech. Rep. Fish. Aquat. Sci.*, No. 1751.

- Cunjak, R.A., Caissie, D., El-Jabi, N., Hardie, P., Conlon, J.H., Pollock, T.H., Giberson, D.J., and Komadina-Douthwright, S. (1993). The Catamaran Brook (New Brunswick) habitat research project: biological, physical and chemical conditions (1990–1992). *Can. Tech. Rep. Fish. Aquat. Sci.* No. 1914.
- Cunjak, R.A. & Therrien, J. (1998) Inter-stage survival of wild juvenile Atlantic salmon, *Salmo salar* L. *Fisheries Management and Ecology*, 5, 209–223.
- Dauphin, G., Prévost, E., Adams, C.E., Boylan, P., (2009). A Bayesian approach to estimating Atlantic salmon fry densities using a rapid sampling technique. *Fish. Manag. Ecol.* 16, 399–408.
- Elliott, J. 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.* Edited by E. Prévost and G. Chaput. INRA Editions, Paris, France. pp. 25–55.
- El-Shaarawi A.H., Zhu R., Joe H. (2011). Modelling species abundance using the Poisson-Tweedie family. *Environmetrics*, 22, 2, 152-164
- Gelman, A., Carlin, J.B., Stern H.S., Rubin, D.B. (2004). **Bayesian Data Analysis. Second Edition.** Chapman & Hall, London, New-York. Col. Texts in Statistical Science, 668 pp.
- Gelman, A., 2006. Prior distributions for variance parameters in hierarchical models. *Bayesian Anal.*, 1, 3, 515–534
- Girard I.L, Grant J.WA, Steingrímsson S.O. (2004). Foraging, growth, and loss rate of young-of-the-year Atlantic salmon (*Salmo salar*) in relation to habitat use in Catamaran Brook, New Brunswick, *Can. J. Fish. Aquat. Sci.*, 61, 12, 2339-2349.
- Hankin, D.G. (1984). Multistage sampling designs in fisheries research: applications in small streams. *Can. J. Fish. Aquat. Sci.* 41, 1575–1591.
- Imre I., Grant J.WA., Cunjak R.A. (2005). Density-dependent growth of young-of-the-year Atlantic salmon *Salmo salar* in Catamaran Brook, New Brunswick, *Journal of Animal Ecology*, 74, 3, 508–516.
- Jensen A.J., Johnsen B. O. (1999), The functional relationship between peak spring floods and survival and growth of juvenile Atlantic Salmon (*Salmo salar*) and Brown Trout (*Salmo trutta*), *Functional ecology*, 13, 778–785
- Jonsson, N., Jonsson, B., & Hansen, L. P. (1998). The relative role of density-dependent and density-independent survival in the life cycle of Atlantic salmon *Salmo salar*. *Journal of Animal Ecology*, 67, 751-762.
- Klemetsen A., Amundsen P.A, Dempson J.B., Jonsson B., N. Jonsson, O'Connell M.F, Mortensen E. (2003), Atlantic salmon *Salmo salar* L., brown trout *Salmo trutta* L. and Arctic charr *Salvelinus alpinus* (L.): a review of aspects of their life histories, *Ecology of Freshwater Fish*, 12, 1, 1–59.

- Limburg, K.E., and J.R. Waldman. (2009). Dramatic declines in North Atlantic diadromous fishes. *Bioscience*, 59, 955-965.
- Lindén, A., & Mäntyniemi, S. (2011). Using the negative binomial distribution to model overdispersion in ecological count data. *Ecology*, 92, 7, 1414-1421.
- Lunn, D., Spiegelhalter, D., Thomas, A., & Best, N. (2009). The BUGS project: Evolution, critique and future direction. *Statist. Med. (Statistics in medicine)*, 28, 3049-3067.
- Mäntyniemi, S., Romakkaniemi, A., Arjas, E. (2005). Bayesian removal estimation of a population size under unequal catchability. *Can. J. Fish. Aquat. Sci.*, 62, 291-300.
- Martin, T. G., Wintle, B. A., Rhodes, J. R., Kuhnert, P. M., Field, S. A., Low-Choy, S. J., Tyre, A. J., et al. (2005). Zero tolerance ecology: improving ecological inference by modelling the source of zero observations. *Ecology Letters*, 8, 11, 1235-1246.
- Michielsens, C.G.J., McAllister, M.K., Kuikka, S., Mäntyniemi, S., Romakkaniemi, A., Pakarinen, T., Karlsson, L., Uusitalo, L. (2008). Combining multiple Bayesian data analyses in a sequential framework for quantitative fisheries stock assessment. *Can. J. Fish. Aquat. Sci.* 65, 962-974
- Mills, D. (1989). **Ecology and Management of Atlantic salmon**. Chapman and Hall, London - New-York, 368pp
- Milner, N.J., Elliott, J.M., Armstrong, J.D., Gardiner, R., Welton, J.S., and Ladle, M. (2003). The natural control of salmon and trout populations in streams. *Fish. Res.* 62, 111–125.
- Mitchell S.C, Cunjak R.A. (2007a). Relationship of upstream migrating adult Atlantic salmon (*Salmo salar*) and stream discharge within Catamaran Brook, New Brunswick, *Can. J. Fish. Aquat. Sci.*, 64, 3, 563-573.
- Mitchell S.C, Cunjak R.A. (2007b). Stream flow and beaver dams: roles in the structuring of stream fish communities within an anadromous salmon dominated stream. *Journal of Animal Ecology*, 76, 6, 1062-1074
- Parrish, D.L., Behnke, R.J., Gepphard, S.R., McCormick, S.D., and Reeves, G.H. 1998. Why aren't there more Atlantic salmon (*Salmo salar*)? *Can. J. Fish. Aquat. Sci.*, 55, Suppl. 1, 281–287.
- Randall, R.G. 1982. Emergence, population densities, and growth of salmon and trout fry in two New Brunswick streams. *Can. J. Zool.* 60, 2239–2244.
- Rivot, E., Prévost, E. (2002). Hierarchical Bayesian analysis of capture-mark-recapture data. *Can. J. Fish. Aquat. Sci.*, 59, 1768-1784.
- Rivot, E., Prévost, E., Cuzol, A., Baglinière, J.L., Parent, E. (2008). Hierarchical Bayesian modelling with habitat and time covariates for estimating riverine fish population size by successive removal method. *Can. J. Fish. Aquat. Sci.*, 65, 117-133.

- Robert, C. P. (1996). **Méthodes de Monte carlo par Chaines de Markov**. Coll. Statistique mathématique et Probabilité (Economica, Paris.). Economica, Paris, 340 pp
- Ruiz, P., and Laplanche, C. (2010). A hierarchical model to estimate the abundance and biomass of salmonids by using removal sampling and biometric data from multiple locations. *Can. J. Fish. Aquat. Sci.*, 67, 12, 2032-2044
- Shearer, W. M. (1992). **The Atlantic salmon Natural history, exploitation, and future management**. Fishing News Books, Osney Mead, Oxford, England, 244 pp.
- Sileshi G., Hailu G., Nyadzi G.I., (2009). Traditional occupancy-abundance models are inadequate for zero-inflated ecological count data. *Ecological Modelling*. 220, 5, 1764-1775
- Steingrímsson, S.Ó. & Grant, J.W.A. (2003) Patterns and correlates of movement and site fidelity in individually tagged young-of-the-year Atlantic salmon (*Salmo salar*). *Canadian Journal of Fisheries and Aquatic Sciences*, 60, 193–202.,
- Taylor B.R, MacInnis C., Floyd T.A. (2010). Influence of rainfall and beaver dams on upstream movement of spawning Atlantic salmon in a restored brook in Nova scotia, Canada. *River research and applications*, 26, 2, 183-193.
- Tetzlaff D., Soulsby C., Youngson A.F, Gibbins C., Bacon P.J, Malcolm I.A., Langan S. (2005). Variability in stream discharge and temperature: a preliminary assessment of the implications for juvenile and spawning Atlantic salmon, *Hydrology and Earth System Sciences*, 9, 3, 193-208
- Vaudor, L., Lamouroux, N., & Olivier, J.-M. (2011). Comparing distribution models for small samples of overdispersed counts of freshwater fish. *Acta Oecologica*, 37, 3, 170-178.
- Webb, J., Verspoor, E., Aubin-horth, N., Romakkaniemi, A., Amiro, P. (2007). **The Atlantic Salmon, Genetics, Conservation and Management**. Verspoor, E., Stradmeyer, L., Nielsen, J.L. eds. Blackwell Publishing Ltd. 520 pp
- Wyatt, R.J. (2002). Estimating riverine fish population size from single- and multiple-pass removal sampling using a hierarchical model. *Can. J. Fish. Aquat. Sci.*, 59, 695-706.
- Wyatt, R.J. (2003). Mapping the abundance of riverine fish population: intergrating hierarchical Bayesian models with a geographic information system (GIS). *Can. J. Fish. Aquat. Sci.*, 60, 997-1006

Appendix I : Probability distributions and functions used in the model

Table 1 : Probability distribution used in the model

Distribution	Notation	Density function
Normal	$X \mu, \sigma^2 \sim \text{Normal}(\mu, \sigma^2)$	$p(X \mu, \sigma^2) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left(-\frac{1}{2\sigma^2}(X - \mu)^2\right)$
Poisson	$X \lambda \sim \text{Poisson}(\lambda)$	$p(X \lambda) = \frac{1}{X!} \lambda^X e^{-\lambda}$ $X = 1, 2, \dots$
Gamma	$X \alpha, \beta \sim \text{Gamma}(\alpha, \beta)$	$p(X \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} X^{\alpha-1} e^{-\beta X}$ $X > 0$
Uniform	$X a, b \sim \text{Uniform}(a, b)$	$p(X a, b) = \frac{1}{b - a}$ $b > X > a$
Beta	$X \alpha, \beta \sim \text{Beta}(\alpha, \beta)$	$p(X \alpha, \beta) = \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha)\Gamma(\beta)} X^{\alpha-1} (1 - X)^{\beta-1}$ $0 > X > 1$
Binomial	$X n, p \sim \text{Binomial}(n, p)$	$p(X n, p) = \binom{n}{X} p^X (1 - p)^{n-X}$ $X = 1, 2, \dots, n$

Table 2 : Functions used in the model

Function	equation
logit	$z(a, h, y, r, k) = \log\left(\frac{p(a, h, y, r, k)}{1 - p(a, h, y, r, k)}\right)$
logit⁻¹	$p(a, h, y, r, k) = \frac{e^{z(a, h, y, r, k)}}{1 + e^{z(a, h, y, r, k)}}$

Appendix II: Winbugs[©] Model

model

```
{
  #Parameters for z

  mu.z~dnorm(0,0.01)
  tau.z<-1/(sd.z*sd.z)
  sd.z~dunif(0,10)

  #Parameters for d

  CV.d2<-1
  CV.d1~dunif(0,5)

  for (A in 1:3)
  {
    for (R in 1:6)
    {
      for (T in 1:21)
      {

        CV.d[A,T,R]<-B[A,T,R]*CV.d1+(1-B[A,T,R])*CV.d2

      }
    }
  }

  #Effect of interaction between HAB/AGE

  for (H in 1:5)
  {
    for (A in 1:3)
    {
      alpha.d[H,A]~dnorm(0,0.01)
      al.d[H,A]<-alpha.d[H,A]-mean(alpha.d[,A])
    }
  }

  #Parameters mixture model

  p.B[1,1]~dbeta(4,1)
  p.B[2,1]~dbeta(4,1)
  p.B[3,1]~dbeta(1,1)
  p.B[4,1]~dbeta(1,1)
  p.B[5,1]~dbeta(1,4)
  p.B[6,1]~dbeta(1,4)
```

```

p.B[1,2]~dbeta(4,1)
p.B[2,2]~dbeta(4,1)
p.B[3,2]~dbeta(1,1)
p.B[4,2]~dbeta(1,1)
p.B[5,2]~dbeta(1,4)
p.B[6,2]~dbeta(1,4)

p.B[1,3]~dbeta(4,1)
p.B[2,3]~dbeta(4,1)
p.B[3,3]~dbeta(1,1)
p.B[4,3]~dbeta(1,1)
p.B[5,3]~dbeta(1,4)
p.B[6,3]~dbeta(1,4)

for (A in 1:3)
{
for (R in 1:6)
{

for (T in 1:21)
{
B[A,T,R]~dbern(p.B[R,A])

}
}
}

#Effect interactions Age/Reach/Year

Bet2.d~dnorm(0,0.1)

for (A in 1:3)
{

mu.bet1[A]~dnorm(0,0.01)
tau.bet1[A]<-1/(sd.betH[A]*sd.betH[A])
sd.bet1[A]~dunif(0,10)

for (R in 1:6)
{

for (T in 1:21)
{
Bet1.d[A,T,R]~dnorm(mu.betH[A],tau.betH[A])

for (H in 1:5)
{
log.E.d[H,A,T,R]<-B[A,T,R]*(bet1.d[A,T,R]+al.d[H,A])+(1-B[A,T,R])*bet2.d

}

}

}

```

```
}  
}  
}
```

#Extrapolation

```
for (T in 1:21)  
{  
  
  for (A in 1:3)  
  {  
  
    for (R in 1:6)  
    {  
  
      for (H in 1:5)  
      {  
  
        lambda'[H,A,T,R]<- E.d'[H,A,T,R]*St[H,R]  
  
        N'[H,A,T,R]~dpois(lambda'[H,A,T,R])  
  
        E.d'[H,A,T,R]<-exp(log.E.d[H,A,T,R])  
  
        }  
        }  
        }  
        }  
  
      for (i in 1:K)  
      {  
  
        #d  
  
        d[i]~dgamma(r.d[i],mu.d[i])  
        mu.d[i]<-1/(E.d[i]*CV.d[a[i],y[i],s[i]]*CV.d[a[i],y[i],s[i]])  
        E.d[i]<-exp(log.E.d1[i])  
        log.E.d1[i]<-log.E.d[h[i],a[i],y[i],s[i]]  
        r.d[i]<-1/(CV.d[a[i],y[i],s[i]]*CV.d[a[i],y[i],s[i]])
```

Posterior Checking

```
d.rep[i]~dgamma(r.d[i],mu.d[i])
```

#z et p

```
p[i]<-exp(z[i])/(1+exp(z[i]))  
z[i]~dnorm(mu.z,tau.z)
```

```
#N

lambda[i]<-d[i]*S[i]
N1[i]~dpois(lambda[i])I(,1000)

N2[i]<-N1[i]-c1[i]
N3[i]<-N2[i]-c2[i]
N4[i]<-N3[i]-c3[i]
N5[i]<-N4[i]-c4[i]

# Likelihood

c1[i]~dbin(p[i],N1[i])
c2[i]~dbin(p[i],N2[i])
c3[i]~dbin(p[i],N3[i])
c4[i]~dbin(p[i],N4[i])
c5[i]~dbin(p[i],N5[i])

}

}
```




Département : Pôle Halieutique
Spécialisation : Sciences Halieutiques et Aquacoles
option : Ressources et écosystèmes aquatiques
Enseignant responsable : M. Didier Gascuel

Auteur(s) : Mlle Clarisse Boulenger

Date de naissance : 11/08/1987

Nb pages : 32 Annexe(s) : 5

Année de soutenance : 2011

Organisme d'accueil : UMR ESE INRA/Agrocampus Ouest
Adresse : 65 rue de Saint Brieu, CS 84215, 35042 Rennes Cedex

Maître de stage : M. Etienne Rivot
Co-Encadrants : M. Richard Cunjak, M. Jean-Marc Roussel, M. Etienne Prevost

Titre français : Variabilité spatio-temporelle de l'abondance de juvéniles de Saumon Atlantique dans un cours d'eau canadien : Approche par Modélisation Bayésienne

Titre anglais : Spatio-temporal variability of Atlantic salmon juvenile abundance in a Canadian stream: A Bayesian Modeling Approach

Résumé :

Un Modèle Hiérarchique Bayésien (MHB) a été construit pour estimer la distribution spatio-temporelle des juvéniles de Saumon Atlantique (*Salmo salar*) dans le Catamaran (Nouveau Brunswick, Canada) grâce à un jeu de données important d'enlèvements successifs par pêche électrique. La partie la plus originale de la modélisation concerne l'utilisation d'un modèle de mélange introduit pour représenter la distribution spatio-temporelle de la densité qui permet de capturer à la fois les densités très élevées et très faibles. De plus ce modèle de mélange introduit une variable latente qui a une interprétation écologique liée à l'accessibilité de certains secteurs du Catamaran par les géniteurs au moment de la reproduction.

Ce travail est une application illustrant comment les MHB sont des outils puissants et flexibles pour l'utilisation de données sur plusieurs sites et années d'enlèvements successifs et ce modèle fournit un cadre pour d'autres recherches et la collecte de données dans le futur.

Le MHB fournit une série temporelle d'estimation d'abondance de 21 années pour les 3 classes d'âge de la phase juvénile. À l'échelle du bassin versant, les séries temporelles d'estimation d'abondance ont montré une importante variabilité interannuelle des juvéniles d'âge 0+. De plus, la série temporelle d'estimation d'abondance des 0+ a permis une première étude de la relation de stock-recrutement. Un contrôle environnemental sur le succès de recrutement par le débit a été mis en évidence. La structure spatialisée du modèle a donné accès à des estimations de la contribution relative des différents secteurs à la production de juvénile. La plus grande variabilité temporelle de contribution entre les secteurs est observée pour les 0+ ce qui semble lié à la variabilité interannuelle de l'accessibilité des secteurs durant la migration des reproducteurs. Une migration vers l'amont des juvéniles du secteur Gorge vers Middle entre les classes d'âge 0+ et 1+ a aussi été mis en évidence.

Abstract :

In this work, a Hierarchical Bayesian Model (HBM) was built to estimate the spatio-temporal distribution of wild Atlantic salmon (*Salmo salar*) juveniles in the Catamaran brook (New Brunswick, Canada) from a large data set of successive removal data via electrofishing. The most original part of our modelling approach is the mixture modelling approach introduced to model the spatio-temporal distribution of the density that enables to capture both very high and very low densities. Moreover, the mixture model introduced a latent variable with an ecological interpretation linked with the process of accessibility to certain sectors of the Catamaran brook.

This work is an application illustrating how the HBM framework is powerful and flexible for processing multiyear and multisite successive removal data and this model provides a framework for structuring further research and data collection in the future.

The HBM provides 21-years time series of abundance estimates for the three age classes of A. salmon juvenile in the Catamaran brook. At the scale of the whole watershed, the time series of abundance estimates highlighted a huge between year variability of the 0+ juveniles. Moreover, the time series of 0+ abundance estimates allowed a first examination of the Stock-Recruitment relationship but only an environmental control of the Recruitment success by the water discharge was pointed out. The spatialized structure of the model gave access to estimates of the relative contribution of each sector to the juvenile production. The greatest temporal variability of the contribution between reaches was observed for the 0+ age group which seems related to the between year variability of accessibility during the spawning migration. An upstream migration of juveniles from the Gorge to the Middle reach between the 0+ and the 1+ stages was also pointed out.

Mots-clés : Modèle Hiérarchique Bayésien, Modèle de mélange, Abondance, Juvéniles, *Salmo salar*, Canada
Key-words : Hierarchical Bayesian Model, Mixture model, Abundance, Juveniles, *Salmo salar*, Canada