Impact of local pollution on fish abundance using geostatistical simulations

Eric de Oliveira, Nicolas Bez

Impact of local pollution on fish abundance using geostatistical simulations

Eric DE OLIVEIRA (1), Nicolas BEZ (1), Etienne PREVOST (2)
(1) Centre de Géostatistique - Fontainebleau - France

(2) Laboratoire d'Ecologie Aquatique - INRA - Rennes - France

Key words: Geostatistics, Simulations, Pollution, Impact, Fish Farm, Salmon.

Abstract:

The objective of this paper is to quantify the impact of human activities on the abundance of fresh water fish populations. In the case of populations of salmon juvenile on the Scorff river (Brittany, France), two punctual pollutions are caused by the discharges of two fish farms. Two perturbed sections have been identified on the main course of the river where salmon juvenile abundance is reduced. Assuming the pollution does not involve any change of density of juveniles salmons in the unperturbed sections, geostatistical simulations have been used to quantify the impact of the fish farms on the abundance of juvenile salmon. To this end, spatial structures of densities are analysed and modelled for two different cases : one using all the data to describe the "actual structure"; one using only the unperturbed data to estimate the "possible structure" of juvenile salmon in the absence of pollution. In each case, the structural model is used to simulate what could be the distribution of juveniles in areas perturbed by the fish farms. The difference between simulations conditioned by all the data or only by the unperturbed data, allows us to quantify the impact of the fish farms in terms of numbers of juveniles lost.

1. INTRODUCTION

Fish farms located along rivers usually derive water from the river to feed rearing facilities and redirect the water to the river after some cleaning. Whatever the quality of the cleaning, the circulation of the water through the farm modifies the physical and the chemical properties of the water and thus modifies the living conditions of wild organisms in the vicinity of the output flow of fish farms. The problem of evaluating the damages caused to wild populations by pollution has become more important as both the production of fish farms and the environmental concerns have increased.

A large number of studies deal with the physiological consequences of different kinds of pollution at an individual scale (Lacroix, 1989; Fuiman, 1993; Kime, 1995; Waters, 1995). At a population scale, quantitative evaluations of the impact of local pollutions are very rare (Rose et al., 1993, proposed an individual based model of the consequences of environmental degradations on the population dynamic of *Morone saxatilis)*. Generally speaking, the main difficulty consists in estimating what could be the status of the population in absence of perturbation.

In this regard, one favourable case is when series of observations made before and after the appearance of a pollutant discharges are available. Another favourable case concerns local pollution: perturbed and unperturbed areas can be differentiated and the characteristics of the fish distribution observed on unperturbed area can be used to simulate what would be the fish distribution in the perturbed areas which is then compared with the actual situation.

In this paper, we deal with the two fish farms (Pont Kerlo and Pont-Callec, Figure 1) located on the main course of the Scorff (Brittany, France). The Scorff is a small river, 75 kilometres long, with mean annual discharge of 5 $m^3.s^{-1}$ and a drainage basin of 480 km^2. Prévost (1999) has shown a statistically significant reduction of the densities of juveniles of Atlantic salmon (*Salmo salar*) downstream from each fish farm. In other words, Prévost has demonstrated that due to the fish farm discharges, some of the wild salmon production in the Scorff river is lost. But still, how much of the production is being lost remains unknown. The present paper addresses the question of the quantification of this loss. It does not intend to test again the presence or the absence of perturbation due to pollution. According to Prévost (op. cite), the river has been split into two perturbed and three unperturbed sections. The objective of the paper is then to simulate what could be the salmon distribution along a river free of pollution and to derive quantification of the impact of pollution by comparison with the actual distribution of salmon.

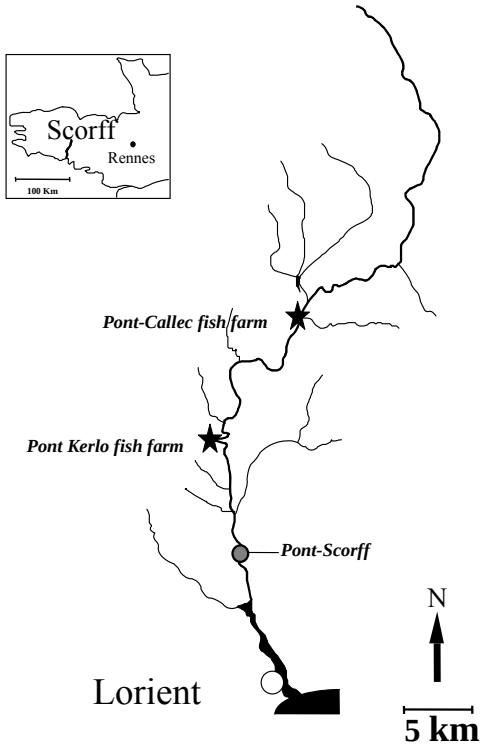

Figure 1. The Scorff river and the location of the 2 fish farms (__) on its main course.

2. MATERIEL AND METHODS

2.1 Sampling

Each autumn (late September – early October) since 1993, the French Institute for Agronomical Research (INRA) survey the Scorff to assess the abundance of young of the year (YOY) salmon. The present study focusses on the two most abundant years, i.e. 1996 and 1997, for illustration purposes. Each year, 39 stations are sampled on the main river, noted S1 to S39. For each station, an index of abundance is obtained by electrofishing (Prévost and Baglinière, 1995). This index is expressed in number of salmon juveniles caught during five minutes of fishing and is proportional to a density (Prévost and Nihouarn, 1999) expressed in number of individual per meter square (ind.m^{-2}). Sampled stations are homogeneous for the physical characteristics of the habitat, shallow areas with coarse bottom substrate and

turbulent flow also called riffles and rapids, i.e. the preferred habitat for YOY salmon in autumn (Baglinière and Champigneulle, 1986; Baglinière et al., 1993).

Habitat types other than riffle/rapids are found in the river, but their use by YOY salmon is lower. By weighting differentially habitat types according to the observed densities collected on several salmon rivers, Prévost and Porcher (1996) quantify the surface area for juvenile salmon production (S_{er}) and express this quantity in a standard unit, the m² of riffle/rapid equivalent. Knowing exhaustively the physical habitat characteristics of the Scorff river (Claude, 1996), a distribution profile of S_{er} along the main course is obtained.

2.2 Sections definition

Prévost (1999) identified 10 perturbed stations (i.e. with fish densities significantly lower than in the other stations), clustered in 2 groups of 5 stations downstream from each fish farm. We then divide the Scorff in 5 sections :

- Section 1: from the estuary to before the first perturbed station (S1 to S4).
- Section 2: the 5 perturbed stations downstream from Pont Kerlo (S5 to S9).
- Section 3: the stations upstream from Pont Kerlo to before the first station perturbed by Pont-Callec fish farm (S9 to S22).
- Section 4: the 5 perturbed stations downstream from Pont-Callec fish farm (S23 to S27).
- Section 5: all the stations upstream from Pont-Callec (S28 to S39).

2.3 Evaluation of the impact of local pollution

If x is a point along the Scorff (distance from the estuary), we note $z(x)$ the salmon juvenile density at point x. Sample locations are noted x_i with $i=1,...,39$ so that $z(x_i)$ and $S_{er}(x_i)$ represent the sample density and the equivalent surface area of production associated to station i. Abundance of salmon juvenile are estimated by multiplying sample densities by their associated surface area of production, so that the abundance of a given section, say Section 2, is defined by :

$$\sum_{i=5}^{9} z(x_i) \times S_{er}(x_i)$$

The quantification of the impact of local pollutions on the abundance of YOY salmon is based on the comparison between the distribution fish could have along the river with pollution and the distribution they could have

without pollution. The former is obtained by a series of 1000 geostatistical simulations conditioned by the complete set of 39 stations (further denoted *sim39)*. The latter is obtained by another series of 1000 simulations where only the 29 unperturbed stations are used for conditioning (further denoted *sim29)*. This amounts to consider two different population sampled at 39 and 29 stations respectively, and to model each density by a random function with its own characteristics (in particular its own variogram) infered from the 39 and 29 respective samples. Conditional simulations of each random function would allow to compare the behaviour of both population on river sections where the first population has been observed at 5 sampling stations, and where the second population has not been sampled at all.

Simulations performed in this work correspond to geostatistical conditional simulations (Matheron, 1973 ; Chilès and Delfiner, 1999) where the conditioning is made by kriging. Once a random function model has been defined, a geostatistical simulation is a numerical model that corresponds to one possible realisation of this random function. A conditional simulation is a simulation that matches the random function characteristics and the data value at samples points. Technically, simulations fall in the scope of spatially consistent Monte-Carlo simulations. They consist in the following steps: gaussian transformation of the data, structural analysis of the gaussian data, conditional simulation and back transformation of the conditional simulations. These steps, with an additional one describing the use of the simulations, are detailed below.

1.1.1 Gaussian transformation of the data

Gaussian transformation comes from the graphical comparison between the cumulated frequencies of the sample data and the corresponding quantiles of a gaussian variable with expectation 0 and variance 1.

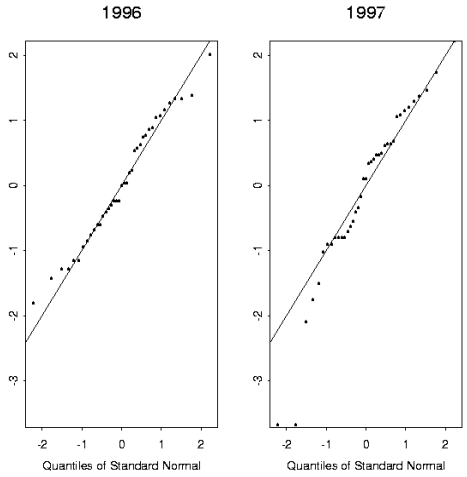

Figure 2. Q-q plot of the gaussian transformed variable (y-axis) versus a gaussian variable (x-axis) with expectation 0 and variance 1.

Graphs have sigmoid shapes, leading to the following transformation :

$$y = \ln\left(\frac{z.\alpha}{0.54 - z}\right)$$

with z the density, y the gaussian transformed density, α a constant controlling the inflexion point, and *0.54* the largest density observed on the Scorff since 1970 (Prévost, 1999). To avoid troubles, the two zero values observed in 1997 have been translated to a small positive value arbitrarily set to *m/100* where *m* is the annual mean density. Graphical comparison between quantiles of the transformed data and of the gaussian variable (Figure 2) are used to test the transformation.

1.1.2 Structural analysis of the gaussian data

Experimental variograms are calculated with either 39 or 29 samples. After several trials, we choose a distance lag of 1500 m (the average distance between two consecutive stations is 1280 m). Once the functions needed for each model are set, estimation of the model parameters is based on a weighted least square procedure where the weights are equal to the number of pairs of points used for the computation of each point of the variogram. Variograms calculated with all the data are modelled by the combination of a nugget effect and a spherical model. Variograms calculated with the

unperturbed samples are modelled by the combinations of a nugget effect and a linear model.

1.1.3 Conditional simulations

Simulations are performed every 100 meters along the Scorff (501 points of simulation). For each model, 1000 simulations are generated.
For a one-dimension (1D) spherical variogram, simulations are performed according to the developments presented in the annex. Basically, the river is divided in regular bins equal to the range of the variogram with uniform random origin. On each bin, a linear segment is implemented with fixed slope randomly positive or negative. To insure normality of the simulated values, several independent simulations are performed and averaged.
Random functions with linear structure are simulated sequentially (Chilès and Delfiner, 1999). From one point to the next, we add a random value following a gaussian distribution with mean equal to zero and variance equal to twice the slope of the linear model times the simulation step (100 meters in this case). Starting from 0, 1500 data are first generated as indicated above. A set of 501 data is then chosen randomly.

To condition the simulations, kriging is performed with a unique neighbourhood including respectively all or only the 29 unperturbed sample values (Chilès and Delfiner, 1999).

1.1.4 Back transformation of the simulations

Simulated values are simply back transformed as follow:

$$z = \left(\frac{0.54 \times e^{y}}{\alpha + e^{y}} \right)$$

1.1.5 Use of simulations for estimation purposes

Three profiles are represented : the profile of the mean of the 1000 simulations (this theoretically converges towards the kriging profile), and the profiles for the quantile 5 and 95% of the simulated data.
Having simulated the density every 100 m, abundance of salmon is obtained by multiplying the densities by S_{er} discretized every 100 m. Abundance are then summed up to compute abundance over the 5 river sections. The impact of pollution will be assessed by comparing the histograms of the 1000 abundance obtained from *sim39* and *sim29* on each of the perturbed sections

(sections 2 and 4). The difference between the means of the two histograms provides an estimate of the loss of salmon production due to pollution.
Simulations are performed separately (independently). However, both are conditioned by 29 common points. So, we can not consider that *sim39* and *sim29* are independent, and we can not derive the variance of the estimation of the impact of the pollution by a simple combination of the histograms. Instead, we use the histograms to indicate what is the probability to have abundance without pollution larger than i) the estimated abundance with pollution and ii) the 95% quantile of the simulated abundance with pollution, that is an upper limit of the abundance that has only 5% chance to be exceeded.

3. RESULTS

Densities fluctuate between 0 and 0.38 ind.m^{-2} (Figure 3). In 1996 and 1997, the density profiles are very similar, with a marked increase between the two fish farms, 17 kilometres from the estuary. Two smaller increases in river sections 1 and 5 are observed.

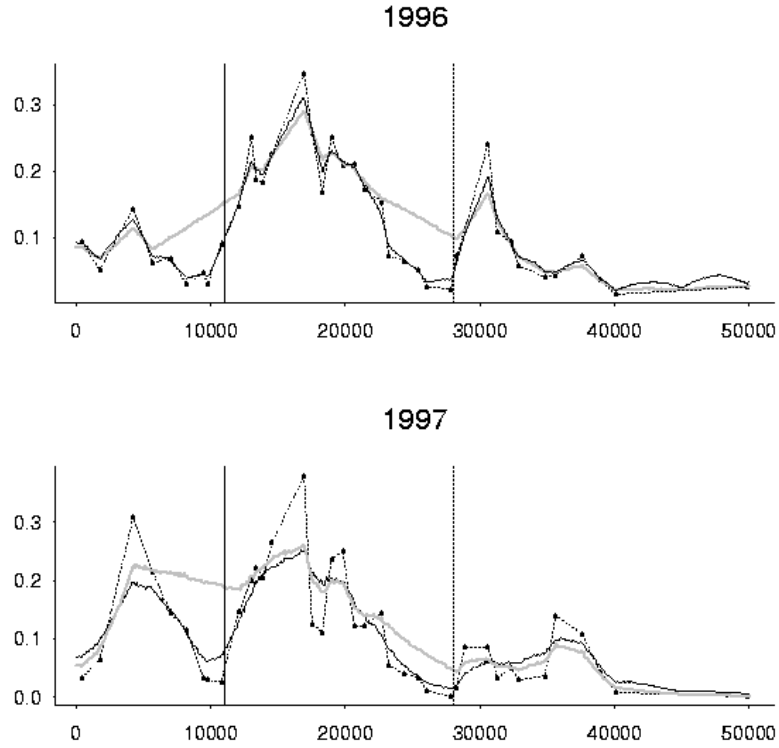

Figure 3. Experimental densities and simulated densities (ind.m^{-2}) along the Scorff. Black dotted line, the experimental profile, black line average profile of the *sim39* and grey line average profile of the *sim29*.Vertical line, the position of the Pont Kerlo fish farm. Vertical dotted line, position of the Pont-Callec fish farm.

Over the four experimental variograms (two years times two sets of data), the nugget effect is small compared to the sill (Figures 4 and 5). Nonetheless nugget effect and variance (variograms's sills) are more important in 1997 than in 1996. Given a set of locations (i.e. 39 or 29 data points), the spatial structures of 1996 and 1997 are similar.

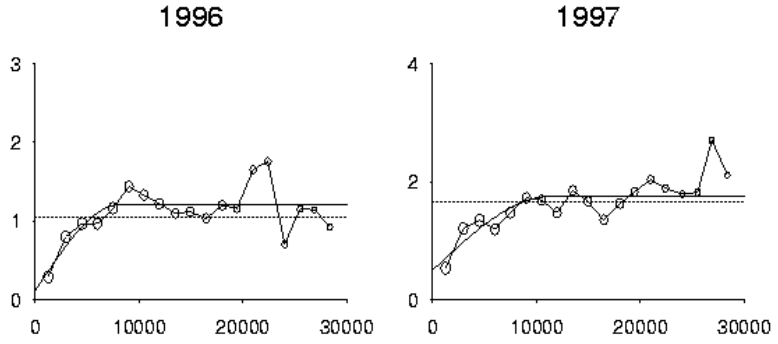

Figure 4. Experimental variogram and model: real case (39 stations). On the x-axis, the distance h (meters), and on y-axis variogram value. Horizontal dotted line, experimental variance.

The nugget effect and the experimental variance are more important in 1997 than in 1996.

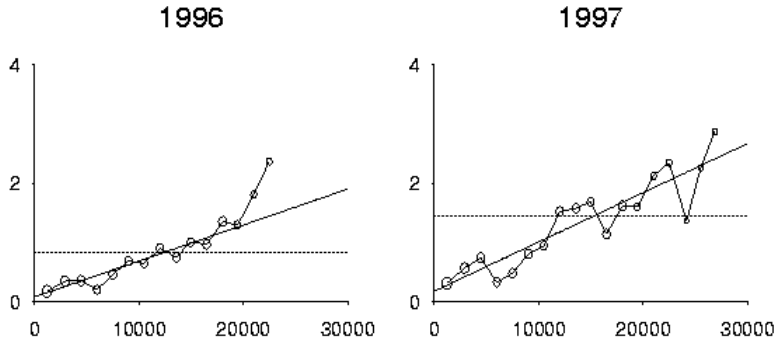

Figure 5. Experimental variogram and model: unperturbed case (29 stations). On x-axis, the distance h (meters), and on y–axis variogram value. Horizontal dotted line, experimental variance.

In 1996, for the three unperturbed sections, the average profile of the simulations respect the experimental profile (Figure 2). The average profile of *sim29* and *sim39* (equivalent to the kriging profile) is smoother than the experimental profile, especially for the 1997 profiles which is due to a larger nugget effect that year. The mean possible profile in case of no pollution is presented Figure 6 with an indication of the fluctuations around this mean profile (90% of the simulations stand between the dotted lines). Annual average abundance of the juvenile population along the river computed with *sim39* is similar to the experimental abundance (Table 1). Using *sim29* leads to an increase of overall abundance of 20% both in 1996 and 1997 (Table 1).

Table 1. Experimental annual abundance for 1996, 1997, annual average and quantile 5 and 95% of abundance obtained from *sim39* and *sim29*.

		Experimental	*sim39*	*sim29*
	quantile 5		10670	14640
1996	mean	21710	21860	26040
	quantile 95		26040	39920
	quantile 5		10120	15950
1997	mean	20080	19740	23660
	quantile 95		26830	29920

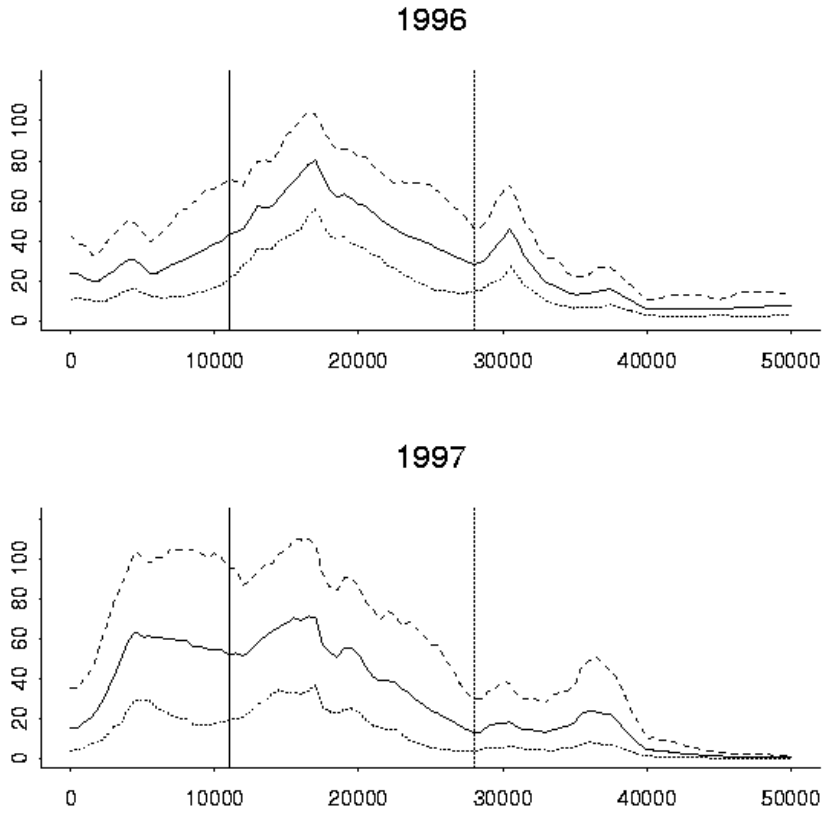

Figure 6. Quantile 5%, 50% and 95% of the sim29 (density in ind.m^{-2}) along the Scorff. Black small dotted line quantile5%, black line quantile 50%, black large dotted line quantile 95%. Vertical line, the position of Pont Kerlo fish farm. Vertical dotted line, position of the Pont-Callec fish farm.

Looking at perturbed sections 2 and 4 responsible for most of the above difference, we can see that the mean abundance is usually two times larger for *sim29* than for *sim39* (Table 2). This goes together with a large increase in the variance, i.e. the range of abundance obtained when conditioning with the unperturbed data is much larger. In the Scorff river, pollution has a negative impact on the salmon population and loss of production estimates are 4100 YOY in 1996 and 3200 YOY in 1997. In 1996, one fish farm appears to have a greater impact than the other.

From *sim39* we know the expected abundance per river section and also the level of abundance that is exceeded only 5 times out of 100 in the actual polluted situation. From *sim29*, we know that, if there was no pollution, the abundance would have:

- at least 90% chance to be larger than the expected one (Table 2, P1),
- 70-85% chance to be larger than the abundance in an optimistic case in 1996 (Table 2, P2),
- and only 50% chance to be larger than the abundance in an optimistic case in 1997 (Table 2, P2). In other words, compared to an optimistic estimation of the abundance of YOY salmon for the polluted river, the predicted abundance in case of no pollution are equally larger or smaller than this abundance.

Table 2. Mean (in number of salmon), standard deviation (σ), and quantile 5% and 95% of simulated abundance for river sections 2 and 4, obtained from *sim39* and *sim29*.
Impact : Estimation of the impact in number of salmon (difference between the means).
P1: Probability to have abundance without pollution larger than the estimated abundance with pollution.
P2 : Probability to have abundance without pollution larger than the quantile 95% of the simulated abundance with pollution.
Experimental abundance : sum of data values times their surface area (see 2.3).

	1996				1997			
	section 2		section 4		section 2		section 4	
	sim39	sim29	sim39	sim29	sim39	sim29	sim39	sim29
mean	844	1669	2163	5468	1264	2616	1558	3410
σ	142	427	383	1380	300	744	497	1347
95 %	1423	3011	3976	10441	2473	4785	3398	9239
5 %	437	639	1208	1892	587	697	405	729
impact	825		3305		1352		1852	
P1	0.97		0.99		0.96		0.91	
P2	0.72		0.86		0.57		0.50	
experimental abundance	713		1855		836		1019	

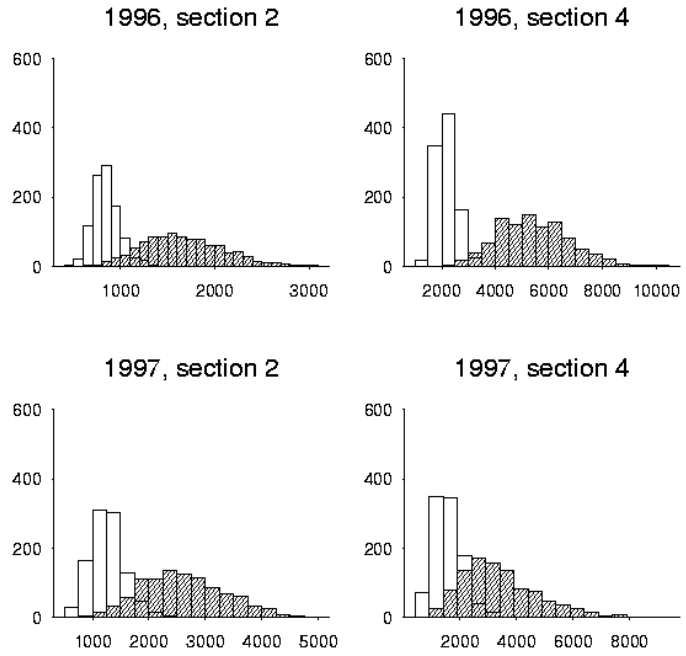

Figure 7. Histograms of abundance in the river sections 2 and 4. In white, the results of the *sim39* and in dashed, the results of the *sim29*.

4. DISCUSSION

This work is based on the results of Prévost (1999) showing that each of the five stations downstream from the fish farms are perturbed. The present work does not intend to discuss these results but rather use them for delineating perturbed and unperturbed sections and estimate loss of production on these river sections. In this context, the possible spatial structure of the YOY salmon in absence of pollution corresponds to the variogram calculated with the 29 unperturbed samples. This implies that small densities observed downstream from fish farms are not compensated by an increase in densities in the unperturbed section. Potential processes that could support compensation are:

– increase of the trophic capacity downstream from the perturbed sections,

– avoidance of the perturbed section by salmon spawners,

– migration of juvenile out of the perturbed areas.

As discussed by De Oliveira et al, in press, none of these three hypotheses hold. The impact of local pollution on YOY salmon population

can be regarded as a loss which can be appraised by the difference between the abundance estimates derived from the two sets of simulations.

The method used in this paper requires identifiable spatial structures, when computed with all the samples but also when computed with the unperturbed samples only. In the present case, 29 stations have been considered as unperturbed which makes the results fragile, i.e. based on a fairly small number of observations. This is also why this study deals only with the two most abundant years (1996 and 1997). Less abundant years do not present easily identifiable spatial structures. Because of the small number of samples available, the parameters of the variogram models and of the gaussian transformations must be interpreted cautiously. A sensibility study of the results to these different parameters could complete this work. We performed another set of simulations replacing the linear variogram by spherical models with large ranges, which did not change the final results.

The average of a large number of simulations is close to kriging and the estimate of salmon population size from the average profile have the meaning of kriging estimates. Because there is no co-localisation between the experimental points and the simulated points and because each variogram model has a nugget component, the average profiles of simulated densities have a smooth aspect. In 1997, when the nugget represents a larger proportion of the variability than in 1996, the smoothing effect is stronger.

The impact of Pont-Callec fish farm is more important than that of Pont Kerlo. This difference is due to the localisation of the two fish farms. The Pont Kerlo fish farm is above a stretch of rivers with a low slope where favourable habitats for YOY salmon are rare. The Pont-Callec facility is above the steepest section of the Scorff with a large potential for YOY salmon production.

Our approach can be adapted to any animal or plant population provided that (i) some areas are perturbed and some are unperturbed with no compensation effect from one to the other, (ii) the sampling covers both the perturbed and the unperturbed areas, and (iii) spatial structures for both areas can be describe from the points of observation.

5. BIBLIOGRAPHY

BAGLINIERE J.L. et CHAMPIGNEULLE A., 1986. Populations estimates of juvenile Atlantic salmon, Salmo salar, as indices of smolt production in the R. Scorff, Brittany. J. Fish. Biol., 29, 467-482.

BAGLINIERE JL., MAISSE G. et NIHOUARN A., 1993. Comparison of two methods of estimating Atlantic salmon, Salmo salar, wild smolt

production. In GIBSON R.J., CUTTING R.E. (eds), Production of juvenile Atlantic salmon, Salmo salar, in natural watters, 189-201, Can. Spec. Publ. Fish. Aquat. Sci., 188.

BARANGE M. et HAMPTON I., 1997. Spatial structure of co-occuring anchovy and sardine populations from acoustic data : implications for survey design. Fish. Oceanogr. 6:2, 94-108.

BEZ N., RIVOIRARD J., GUIBLIN P. et WALSH M., 1997. Covariogram and related tools for structural analysis of fish survey data. E.Y. Baafi and N.A. Schofield (Eds), Geostatistics Wollogong'96, Volume 2. Kluwer, Dordrecht. 1316-1327 p.

CHILES J.P. et DELFINER P., 1999. Geostatistics, Modelling Spatial Uncertainty. Willey edition.

CLAUDE A., 1996. Le recrutement chez le saumon atlantique (Salmo salar L.) dans le Massif Armoricain. Quantifacation des surfaces d'habitat favorables aux juvéniles et estimation de la survie embryo-larvaire sur le Scorff (Morbihan) et l'Oir, affluent de la Sélune (Manche). Mémoire C.E.S.A. opt. Halieutique, E.N.S.A. Rennes, 44p.

CONAN G.Y., 1985. Assessment of shellfish stocks by geostatistcal techniques. ICES C.M. 1985/K:24, 19p.

DE OLIVEIRA E., BEZ N., PREVOST E., in press. Les espaces de l'halieutique. IRD editions.

FOOTE K.G., 1993. Abundance estimation of herring hibernating in a fjord. ICES C.M. 1993/D:45. Statistics Committee.

FUIMAN L.A. (éditeur), 1993. Water quality and the early life stages of fishes. Am. Fish. Soc. Symposium, 14, 172 p.

KIME D.E., 1995. The effects of pollution on reproduction in fish. Rev. Fish Biol. Fish., 5: 52-96.

LACROIX G.L., 1989. Ecological an physiological responses of Atlantic salmon in acidic rivers of Nova Scotia. Water Air Soil Pollution, 46 :375-386.

MATHERON G., 1965. Les variables régionalisées et leur estimation : une application de la théorie des fonctions aléatoires aux sciences de la nature. Masson et Cie, Éditeur, Paris. 306p.

MATHERON G., 1973. The intrinsic random functions, and their applications. Adv. In Appl. Prob., 5: 439-468.

PELLETIER D. et PARMA A.N., 1994. Spatial Distribution of Pacific Halibut (Hippoglossus stenolepis) : An Application of Geostatistics to Longline Survey Data Can. J. Fish. Aquqt. Sci. 51 : 1506-1518.

PETITGAS P., 1993. Geostatistics for Fish Stock assessments: a review and an acoustic application. ICES J. mar. Sci., 50:285-298.

PREVOST E., 1999. Utilisation d'un test de randomisation pour détecter l'effet de rejets polluants dans un cours d'eau : application à l'impact

d'effluents de pisciculture sur la production de juvéniles de saumon atlantique. Bull. Fr. Pêche Piscic., 355 : 369-386.

PREVOST E. et BAGLINIERE J.L., 1995. Présentation et premiers éléments de mise au point d'une méthode simple d'évaluation du recrutement en juvéniles de saumon atlantique (Salmo salar) de l'année en eau courante. In GASCUEL D., DURAND J.L., FONTENEAU A. (eds), Les recherches françaises en évaluation quantitative et modélisation des ressources halieutiques. Actes du colloque, Rennes du 29 juin au 1er juillet 1993, 39-48, ORSTOM Editions, Paris.

PREVOST E. et NIHOUARN A., 1999. Relation entre indicateur d'abondance de type CPUE et estimation de densité par enlèvements successifs pour les juvéniles de saumon atlantique (Salmo salar L.) de l'année. Bull. Fr. Pêche Piscic., 352 :19-30.

PREVOST E. et PORCHER J-P., 1996. Méthodologie d'élaboration de totaux autorisés de captures (TAC) pour le Saumon atlantique (Salmo salar L.) dans le Massif Armoricain. Propositions et recommandations scientifiques. GRISAM, Evaluation et gestion des stocks de poissons migrateurs, Doc. Sci. Tech. 1, 18p.

RIVOIRARD J. et BEZ N., 1997. A 1D Geostatistical Analysis on Norwegian Spring-Spawning Herring A coustic Data in Ofotfjord (December 1994). ICES C.M. 1997/Y:12. 7p.

ROSE K.A., COWAN J.H., HOUDE E.D. et COUTANT C.C., 1993. Individual-based modelling of environmental quality effects on early stages of fishes: a case study using striped bass in FUIMAN L..A. (éditeur), 1993. Water quality and the early life stages of fishes, 125-145, Am. Fish. Soc. Symposium 14.

SAMB B. et PETIGAS P., 1997. Estimation de la précision des campagnes acoustiques au Sénégal par la méthode géostatistique transitive à une dimension. Aquat. Living Resourc., 1997, 10:75-82.

SIMMONDS E.J. et FRYER R.J. 1996. Which are better, random or systematic acoustic surveys? A simulation using North Sea herring as an example. ICES Journal of Marine Science, 53:39-50

WATERS T.F., 1995. Sediment in streams. Sources, biological effects and control. Am. Fish. Soc. Monograph, 7, 251 p.

6. ANNEX

This annex explains how to simulate in 1 dimension a gaussian random function with a spherical variogram.

An origin X_0 having been chosen at random uniformly on (0,a), the straight line is divided into regular bins of length a with subdivision points $X_0 + ka$ where k is an integer value.

$$X_0 \equiv U[0,a]$$

Let ε be a random variable taking +/– $\frac{b}{a}$ with equal probability.

$$E(\varepsilon) = 0$$
$$Var(\varepsilon) = E(\varepsilon^2) = \frac{b^2}{a^2}$$

If we consider the variable X as:

$$X = x - X_0 - ka \equiv U[0,a]$$
$$E(X) = \frac{a}{2} \quad and \quad E(X^2) = \frac{a^2}{3}$$

and Y(x) the random function defined by:

$$Y(x) = \varepsilon.(x - X_0 - ka)$$
$$E(Y(x)) = 0$$
$$Var(Y(x)) = E(\varepsilon^2).E(X^2) = \frac{b^2}{a^2}.\frac{a^2}{3} = \frac{b^2}{3} = c_0$$

The probability that x and x+h belong to the same bin is 1-|h|/a (Matheron, 1973, p. 100, exercise 1).

If h>0:

$$Cov(Y(x),Y(x+h)) = \left(1-\frac{h}{a}\right).E(\varepsilon^2).E(X.(X+h))$$

$$with \qquad X \equiv U[0,a-h]$$

$$Cov(Y(x),Y(x+h)) = \left(1-\frac{h}{a}\right).\frac{b^2}{a^2}.\int_0^{a-h}\frac{X^2+hX}{a-h}dX = \frac{b^2}{a^3}\left(\frac{(a-h)^3}{3}+\frac{h(a-h)^2}{2}\right)$$

$$Cov(Y(x),Y(x+h)) = c_0\left(1+\frac{h^3}{2a^3}-\frac{3h}{2a}\right)$$

If h<0

$$Cov(Y(x),Y(x+h)) = \left(1+\frac{h}{a}\right).E(\varepsilon^2).E(X.(X+h))$$

$$with \qquad X \equiv U[-h,a]$$

$$Cov(Y(x),Y(x+h)) = \left(1+\frac{h}{a}\right).\frac{b^2}{a^2}.\int_{-h}^{a}\frac{X^2+hX}{a+h}dX = \frac{b^2}{a^3}\left(\frac{a^3}{3}+\frac{ha^2}{2}-\frac{h^3}{6}\right)$$

$$Cov(Y(x),Y(x+h)) = c_0.\left(1+\frac{|h|^3}{2a^3}-\frac{3|h|}{2a}\right)$$

which in both cases is nothing but the spherical function.

To get a gaussian random function, we consider the average of N such independent random functions:

$\frac{1}{N}\sum_{i=1}^{N}Y_i(x)$ converges towards a gaussian variable $(0,\frac{b}{\sqrt{3.N}})$