Fisheries and Oceans Canada | Pêches et Océans Canada
Science | Sciences

Canadian Science Advisory Secretariat (CSAS)

Research Document 2013/113

Gulf Region

Methods and models used in the 2012 assessment of the snow crab (*Chionoecetes opilio*), stock in the southern Gulf of St-Lawrence

E. Wade, M. Moriyasu, and M. Hébert

Fisheries and Oceans Canada
Gulf Region
P.O. Box 5030
Moncton, New Brunswick
E1C 9B6

March 2014

Foreword

This series documents the scientific basis for the evaluation of aquatic resources and ecosystems in Canada. As such, it addresses the issues of the day in the time frames required and the documents it contains are not intended as definitive statements on the subjects addressed but rather as progress reports on ongoing investigations.

Research documents are produced in the official language in which they are provided to the Secretariat.

Published by:

Fisheries and Oceans Canada
Canadian Science Advisory Secretariat
200 Kent Street
Ottawa ON K1A 0E6

http://www.dfo-mpo.gc.ca/csas-sccs/
csas-sccs@dfo-mpo.gc.ca

ISSN 1919-5044

Correct citation for this publication:

Wade, E., Moriyasu, M., and Hébert, M. 2014. Methods and models used in the 2012 assessment of the snow crab (*Chionoecetes opilio*), stock in the southern Gulf of St-Lawrence. DFO Can. Sci. Advis. Sec. Res. Doc. 2013/113. v + 50 p.

TABLE OF CONTENTS

ABSTRACT

This document describes the methods and procedures specific to the 2012 assessment of the snow crab stock of the southern Gulf of St-Lawrence. Changes to methodologies recommended during the framework snow crab assessment methodologies in 2005 and 2011 were adopted. The abundances in number or weight of snow crab stages are estimated using a geostatistics method called kriging with external drift. Estimates are made for the entire southern Gulf snow crab biological unit and for each of the four snow crab fishing areas. An inference and forecast Bayesian model is used to predict the incoming recruitment of commercial-sized adult male crab which will be available for the next year's fishery. The biomass estimates and the forecast model are used for the risk analysis of catch options for the snow crab fishery for the southern Gulf of St. Lawrence overall. An independent analysis of the 2012 data and kriging estimates are provided as an appendix to the document.

Méthodes et modèles servant à l'évaluation de 2012 du stock de crabe des neiges (*Chionoecetes opilio*) dans le sud du golfe du Saint-Laurent

RESUME

Dans le présent document, on décrit les méthodes et les procédures utilisées pour l'évaluation de 2012 du stock de crabe des neiges du sud du golfe du Saint-Laurent. Les modifications de méthodologies recommandées suite aux revues cadres des méthodologies pour l'évaluation du crabe des neiges entreprises en 2005 et en 2011 ont été utilisées. Les abondances, en nombre et en poids par stade de crabe des neiges, sont évaluées dans une approche de géostatistique en utilisant le krigage avec dérive externe. Les estimations d'abondance sont fournies pour la zone du sud du golfe du Saint-Laurent représentant l'unité biologique du crabe des neiges, ainsi que pour les quatre zones de gestion du crabe dans le sud du golfe. Un modèle d'inférence et de prévision Bayésien sert à prédire l'abondance du recrutement de crabe mâle adulte de taille commerciale attendu à la pêche de la prochaine année. Les estimations de biomasses et les prévisions fournies par le modèle servent dans l'analyse de risque des options de captures pour la pêche de crabe des neiges dans l'ensemble du sud du golfe du Saint-Laurent. Une évaluation indépendante des données de 2012 avec krigage est fournie en annexe dans ce document.

INTRODUCTION

This document describes the methods and procedures specific to the 2012 assessment of the snow crab stock of the southern Gulf of St. Lawrence.

During the past two decades of the snow crab survey in the southern Gulf of St. Lawrence (sGSL), adjustments have been introduced to improve the procedure and accuracy of the snow crab assessment. Historically, both the survey area and the spatial density of sampling stations varied through time. Given the various changes that have occurred during the history of the survey, it is a challenge to establish a standard method to compare the complete time series data from 1988 to 2012. Changes to methodologies recommended during the internal reviews of snow crab assessment methodologies in 2005 and 2012 were adopted (DFO 2006, 2012). The most important change resulting from the framework review of 2011 (DFO 2012) is the expansion of the reference area used during the analysis (called kriging polygon) to cover the 20 to 200 fathoms corresponding to the extent of bottom temperature which are favorable for snow crab to cover the SGSL biological unit (Fig. 1).

Analyses have been standardized back to 1997 using a geostatistical method called kriging with external drift technique (KED) along with a consistent study area (kriging polygon). For the 2011 snow crab assessment, the trawl survey was conducted using the procedures defined in the Assessment Framework Workshop of 2005 (DFO 2006), while the survey data were processed according to the recommendations from the 2011 Snow Crab Assessment Methods Framework Science Review (DFO 2012). For the 2011 assessment, all previous year biomass estimates from 1997 to 2010 were re-analyzed using the advised methods (DFO 2012).

In 2012, the sampling grid for the survey was redesigned according to the recommendations from the framework methods review (DFO 2012). The sampling design used equidistant square grids to distribute the samples rather than 10 minute by 10 minute grids (Surette and Wade 2012).

MATERIAL AND METHODS

For the 2012 survey and analysis, the following procedures and parameters were used:

- a survey area with 325 full or partial grids
- a target sampling intensity of 325 stations
- stations randomly assigned to grids in proportion to their area
- data for non commercial crab are densities (number per km^2)
- data for commercial crab derived from size to weight conversion are weights at station (weight per km^2)
- an annual variogram using all data with the exception of one (1) outlier
- a variogram using 25 lags with lag distance of 3 km
- average of global variograms is calculated over three years
- spherical variogram models are used
- kriging with external drift (KED) with depth as covariate using all data
- a regular grid of 100 by 100 for interpolation of densities
- a maximum 8 samples per quadrant are used in neighborhood search (maximum of 32), and
- a new kriging polygon of 57,840 km2 covering the area between 20 to 200 fathoms.

Specific details are described below.

HABITAT

Bathymetry maps of the sGSL (Fig. 2) shows that the area has several banks (Bradelle, American, Orphan) alongside several valleys (Western and Eastern Bradelle, Shediac) and a few deep troughs (Cape Breton and Chaleur). The depth for the sGSL area covered by the snow crab trawl survey ranges from 20 fathom to 200 fathom, with the northern parts of the sGSL showing the deepest waters in the area near the Laurentian Channel. The snow crab stock is therefore bounded away from the shoreline by the warm, shallow waters on the south and west (shores of the Gaspe Peninsula, NB, PEI and Cape Breton Island) and to the north east by the warmer, deep waters of the Laurentian Channel.

Crab distribution versus temperature and depth

In the sGSL, there is a relationship between bottom temperature and depth (Fig.3). In the summer, bottom temperatures in depths between 60 to 120 m tend to be near zero. In deep waters, temperature values are stable near 5 Deg. C and in shallow waters, temperatures increase. Commercial snow crab density tends to be maximal in areas of low bottom temperature (Fig. 4). There is a constant trend of decreasing crab density until bottom temperature reaches 5 Deg. C, where minimal densities occur. Crab densities tend to be lower in shallow water and very deep water. Median crab densities were found to be near zero at depths less than 50 meters and were very low in depths greater than 180 meters (Fig. 5).

SAMPLING

According to the most recent methods framework review of 2011 (DFO 2012), the sampling design was modified. The survey area starting in 2012 was partitioned into equidistant grids (13.36 x 13.36 km) (Fig. 6).These grids cover depths of 20 to 200 fathoms. There were 325 sampling points for the 2012 sampling season with one sample per grid (Surette and Wade 2012).

The grid array was constructed within a Universal Transverse Mercator (UTM NAD83, zone 20) projection. The zone 20 datum encloses the entire survey area. Latitude-longitude coordinates were transformed to and from this space using the rgdal (Keitt et al. 2012) package in R (R Core Team 2012).

Two alternate sampling stations were also generated within each selected cell within the survey area. These alternates are to be used if the bottom proves to be untrawlable at the primary sampling station.

In 2012, a total of 321 successful tows were completed (Fig. 7; Landry et al. 2013).

THE MODELLING APPROACH (GEOSTATISTICS)

The methodology used in the assessment considers crabs per tow divided by the swept area as crab density or crab weight per km^2 for each station and used geostatistics as described below to estimate the total abundance for the estimation area. These values were used in population models and to estimate populations of various categories of crabs (males, females, recruitment, etc.).

Details on the operation of the trawl, biological characterization of snow crab, data collected, and modifications through time are described by Moriyasu et al. (2008). Key protocol modifications were boats, the expansion of the survey area and modification of the swept area estimations during the early nineties.

The data of interest in the analyses described below include position of the station, grid occupied by the station, swept area at the station, abundance (number and weight) by crab category, depth at station and temperature at station.

Prior to analysis, the tow swept areas were calculated from data gathered by acoustic trawl monitoring sensors (Moriyasu et al. 2008).

Conversion of crab densities to weight per km^2

In the analysis of commercial crab categories, estimates of biomass were derived based on the weight of crab sampled at each station. The weight of crab in the catch at each station was calculated as:

$$Wt_k = \sum_{cw=95}^{max.cw} N_{cw.k} * \alpha * CW_k^{\beta} \qquad (1)$$

with Wt_k = total weight of crab (g) at station k, and

$N_{CW.k}$ = catch (in number) of snow crab of carapace width (CW) >= 95 mm at station k.

$\alpha * CW_k^{\beta}$ is the weight (g) to carapace width (mm) relationship with α = 0.0002665 and β = 3.098 (Hebert et al. 2002) used through the time series.

It was assumed that crabs with soft shell conditions observed during the survey will be hard prior to the fishery the following spring. Also, no adjustments to the weight of a crab were made when missing legs were noted.

Distance conversion

For all geostatistical analyses described below, latitude and longitude coordinates were converted to distance from a reference point and corrected for curvature of the earth using the great circle distance formula using decimal degrees:

$$d = r * \mathrm{acos}(\sin(lat1) * \sin(lat2) + \cos(lat1) * \cos(lat2) * \cos(lon2 - lon1)] \qquad (2)$$

where $r = 6378.7$.

Variograms

One of the tools available to assess and summarize spatial autocorrelation is the variogram. The experimental variogram, γ_e (h), provides a description of how the data are related to the distance. It was originally defined by Matheron (1963) as:

$$\gamma_e(h) = \frac{1}{2N(h)} \sum_{(i,j)|h_{i,j} \approx h} (Z(x_i) - Z(x_j))^2 \qquad (3)$$

where N(*h*) is the number of all pairwise data points separated by distance *h* in the dataset and Z(x_i) represents the stochastic process at location x_i. The term semivariogram is also frequently used to designate this relationship (equation 3). A more complete description of variogram theory and analysis methods can be found in Armstrong et al. (1992) and Rivoirard et al. (2000).

The most common model used in fitting the variogram during analysis of the sGSL snow stocks is the spherical model (equation 4 below) (Fig. 8).

$$\gamma(h) = C_0 + C_1 \left[\frac{3h}{2a} - \frac{h^3}{2a^2}\right] \therefore h \leq a \qquad (4)$$

where *a*, C_0 and C_1 are parameters.

Fitting of a theoretical variogram model to an empirical variogram is done by non-linear regression. The variogram fitted can be used to predict the abundance of our variable at locations not sampled (point estimation by kriging) or over a user-defined region (block estimation by kriging).

Annual variograms for the commercial crab density over the entire field from 2006 to 2011 are shown in Figure 8.

For the 2012 analysis, an outlier located north near PEI (Fig. 10) affected the fitting of the variogram such that no apparent autocorrelation was detected when using weighted least-squared fitting. Normal procedure is to remove the outlier during the variogram fitting process but include it in all the other steps of the analysis (Cressie 1993). Appropriate removal of outliers reveals latent spatial dependence and patterns (Rossi et al. 1992). The resulting variogram model once the outlier is removed is shown in Figure 8.

Since the empirical semivariogram is created by averaging the squared differences between pairs of points that are approximately the same distance apart, an outlier can heavily influence this average. Outliers are almost always problematic for kriging, but they are particularly bad when the outliers are scattered randomly throughout the study region (rather than being clustered together). This is because randomly scattered outliers will affect the empirical semivariances at small distances (because they might be right next to low values as in this case), but if the outliers are clustered, the squared difference between two outliers might still be small, allowing for accurate semivariogram estimation at small distances (which is the most important part of the semivariogram because closer neighbors get the highest weights).

3-year Average variogram

Approaches to the temporal stabilizations of variograms were explored during the 2005 and 2011 reviews using various re-weighting and windowing schemes (1 to 10 year moving windows and uniform/linear/exponential decay schemes). A three-year moving window was observed to provide sufficient stabilization for the sGSL snow crab historical data (DFO 2006).

This is done by scaling the empirical variograms by variance before being averaged over a three year period (DFO 2006). Specifically, the averaging is performed as follows:

1. Computing the empirical variograms for the years under consideration labeled as:

   vgm[t], vgm[t-1], vgm[t-2]

   where t = year and each vgm is a vector of the semivariance at each bin.

2. Care is taken so that the size of bins (distances) of the empirical variograms are the same across years (to allow simple averaging).

3. Compute the (total) variance of the data for the years of interest labeled as:

   var[t], var[t-1], var[t-2]

4. Divide each of the empirical variograms by their associated variances to obtain standardized variograms:

   std.vgm[t] = vgm[t]/var[t],

   std.vgm[t-1] = vgm[t-1]/var[t-1],

   std.vgm[t-2] = vgm[t-2]/var[t-2].

5. Compute the average standardized variogram. In the simple case of a three-year lagged mean:

   avg.std.vgm = (std.vgm[t] + std.vgm[t-1] + std.vgm[t-2]) / 3

6. Rescale the average standardized variogram to the variance of the current year Which is then used for kriging calculations:

   final.vgm[t] = avg.std.vgm * var[t]

Variogram models for the commercial crab category averaged over 3 years from 2005 to 2011 are shown in Figure 9.

Kriging with external drift

Kriging with external drift (KED) is a spatial interpolation technique that combines a dependant variable with auxiliary variables (such as depth, temperature, slope, sediment, predator density, etc.) with kriging of the regression residuals. KED is described fully by (Webster and Olivier 2001).

A by-product of Kriging is the Kriging variance (or its square root, the standard error). It is a function of:

(i) the form of spatial variation in the data (modelled, for example, by the variogram) and

(ii) the spatial configuration of the observations in relation to each other and to the estimate.

The KED variance is calculated as the weighted average of the covariances from the new point (s_0) to all calibration points (s_1,...,s_n), plus the Lagrange multipliers (Webster and Olivier 2001):

$$\sigma^2_{KED} = (C_0 + C_1) - C_0^{KEDT} * \lambda_0^{KED} \qquad (5)$$

Goovaerts (1997) describes fully the methods used in estimating the kriging variance when using KED.

Point kriging is then used to estimate density at precise locations. The point estimation process is useful to produce a map of the resource and estimating the average density within a given polygon. By setting a grid of cells over the study area we can estimate the density of snow crab at each node (or cell) and color-code the abundance values to obtain a map of the resource. In this sense, point kriging is an interpolation method with the advantage that it yields a precision index for each estimate on the map. In the snow crab assessment, the kriging grid used to generate the map is defined as a 100 by 100 matrix with a starting reference of 45.5º N. and 66º W in the lower left corner of the grid.

Kriging with external drift using depth

One of the recommendations from the 2005 and the 2011 framework reviews (DFO 2006, 2012) related specifically to the kriging analysis, in which it was suggested that a secondary variable, depth for which the structure is better known, be considered during the analysis and the results compared with other techniques. This was recommended for several reasons but most importantly, because a relationship often exists locally, especially along the extreme edges of crab habitat, between depth and crab density.

The choice of water depth as a predictor of local snow crab density is less based on its direct relevance to local density, than its availability and its correlation with a number of important environmental variables, such as bottom temperature, salinity and bottom type. These variables may directly or indirectly impact local abundance by affecting larval settlement, recruitment, mortality or movement. (Surette et al. 2007).

For the snow crab assessments, a high resolution depth matrix covering the entire Gulf of St. Lawrence was used (Dutil 2011).

Cross validation

In cross validation diagnostics, one sample is removed from the dataset, and the value in its location is predicted using information from the remaining observations. Then the same procedure is applied to the second, and third, and so on to the last sample in the database. Comparison of the average difference between predicted and observed values is made.

The cross validation diagnostics map for the 2012 adult male category is shown in Figure 11.

Cross validation is a general statistical method that may also be used for evaluating other aspects of a geostatistical model, such as the adequacy of competing variogram models, underlying assumptions, structural aspects such as isotropy or anisotropy (Delhomme 1978), and the size of the local neighborhood to be used in kriging.

Surette et al. (2007) showed by using cross validation that the kriging with external drift (KED) technique with depth as a secondary variable performed better overall than ordinary kriging (OK) for the snow crab assessment in the sGSL. Most of the differences between the two techniques occur along the edges of the sampling zone since local relationships between depth and crab density tend to occur most often at these locations.

From the 2005 methods workshop (DFO, 2006), it was concluded that KED was the preferred method.

Local Neighborhood

Rather than kriging values at unknown locations by considering the global sample, we may further relax the stationarity assumption by considering local neighborhoods of sample points. Thus, it is only assumed that the mean is constant, or follows a linear function in the KED case, within a local neighborhood rather than globally. Better predictions are obtained in an analogous manner to a series of linear models approximating a complex function. In cases where there is no relationship between crab density and depth, the KED system will converge towards the OK solution.

The approach adopted here is to use a maximum of 32 nearest neighbors for each interpolated point, with a maximum of 8 per quadrant. The latter constraint limits the impact that points in a given direction may exert, especially along the edges of the study area. Prior testing using cross-validation shows that the optimum number of local neighbors in our case usually lies between 20 and 40, with values within this range showing little quantitative difference. Because a drift model is implicitly fitted to the data when using KED, the local neighborhood must contain a sufficient number of data points to ensure a stable estimate of the drift function parameters and avoid degenerate configurations which would lead to a singular kriging matrix. Thus for consistency, a neighborhood of 32 samples was selected and adopted.

Kriging Polygons

The study area (the area in which the biomass estimates are derived and often called the kriging polygon) is a critical factor in the present assessment method. The area adopted during the most recent review in 2011 for the sGSL assessments is a polygon of 57,840 km^2 which envelops the 20 to 200 fathom depth contours (Fig. 12).

Software

All geostatistical analyses of the sGSL snow crab stocks since the late 1990's were performed using the MATLAB© interpreter language.

For practical purposes, a series of R software scripts were developed to perform the same types of analysis used in this assessment. The gstat library (Pebesma 2004) within R approximates closely the capabilities of the MATLAB© tools.

Estimates of biomass and density contour maps

Following the framework review of 2011 (DFO 2012), population estimates were re-evaluated using the recommended method by applying KED using depth as a covariate and a global variogram using all the available data that was then averaged over 3 years. A neighbourhood of 32 samples was used in order to better linearise the density versus depth relationship if one exists. Data used was the commercial crab weight (weight of crabs per km^2) at each station. Point kriging was used for estimating the population over the sGSL wide polygon of 57,840 km^2. Point kriging was also used to produce the interpolation over a 100 by 100 grid which was then use to create contour maps.

Biomass estimates range from 30,920 t in 2009 to 103,000 t in 2004 (Table 1). Density contour plots of commercial males from 2006 to 2011 using KED were produced (Fig. 13). Crab densities tend to generally decrease towards the edges of the sampling area. These areas are also most affected by fluctuations in crab population due to expansion or contraction of the snow crab range.

A second contour map is produced showing the kriging variance map. It shows the expected error for any location on a map. For example, a kriging variance map for the 2012 adult male category was produced (Fig. 15), which showed lowest variance where samples occur, and highest in unsampled areas.

ESTIMATION OF BIOMASS WITHIN THE MANAGEMENT AREAS

BIOMASS ESTIMATE WITHIN EACH ZONE

Biomass for each snow crab fishing zone was estimated using the method for the global zone but using the polygon specific to the snow crab fishing zone (Fig. 15). The variogram model used is unmodified from that used for the global zone. All samples were used in the geostatistical analysis and the parameters for neighbourhood search (32 samples) and grid definition also were unmodified. The individual zones were defined by polygons shown in Figure 15. Specifically, the kriging with external drift geostatistical method using depth as a covariate was redone for each zone. There are large confidence intervals for the estimates of biomass of the smallest snow crab fishing zones where few samples occurred and which are near the edge of the survey (Table 2).

The surface areas of the polygons used in the assessment are (Fig. 15):

- Zone 12 : 48,028 km^2,
- Zone 19 : 3,833 km^2,
- Zone E : 2,443 km^2, and
- Zone F : 2,438 km^2.

Closed zones and buffers

A zone closed to fishing in Area 12 but which are included in the estimation of crab biomass for Area 12 and for the overall southern Gulf estimates consists of an area around the Irving Whale (Fig. 15). This closed zone contains a surface area of 97 km^2, which represents 0.1% of the southern Gulf polygon of 57,840 km^2.

The specific boundaries of this area closed to fishing are:

- That portion of Crab Fishing Area No. 12 enclosed by straight lines joining the following points in the order in which they are listed:

Point	North Latitude	West Longitude
1	47° 19'30" N	63° 23'36" W
2	47° 25'00" N	63° 23'36" W
3	47° 25'00" N	63° 16'00" W
4	47° 19'30" N	63° 16'00" W
5	47° 19'30" N	63° 23'36" W
Ref: Gulf Fisheries Management Region Close Time Variation Order 2000-013		

Narrow buffer zones are located on the boundaries of Area 19 (Fig. 15). These areas are closed to fishing and are not included in the estimation of crab biomass for Area 12 but are included in the Southern Gulf estimate. The two buffer zones contain a total surface area that represents 0.7% of the southern Gulf polygon of 57,840 km^2 (Figure 12). The buffer zones are described as follows:

- That portion of Crab Fishing Area 12 enclosed by straight lines joining the following points in the order which they are listed:

Buffer zone	Point	North Latitude	West Longitude
1	1	47° 30' 00" N	60° 43' 20" W
	2	47° 32' 12" N	60° 42' 15" W
	3	47° 18' 30" N	60° 18' 00" W
	4	47° 16' 25" N	60° 17' 40" W
	5	47° 30' 00" N	60° 43' 20" W
2	1	46° 21' 40" N	61° 11' 09" W
	2	46° 33' 15" N	61° 34' 12" W
	3	46° 37' 30" N	61° 30' 15" W
	4	46° 25' 40" N	61° 07' 00" W
	5	46° 21' 40" N	61° 11' 09" W
When the geographic boundary of an area is expressed in Latitude and Longitude those point references are based on the geodesic system North American Datum 1983 (NAD83). (Ref:Gulf Region Close Time Variation Order, 2010-039)			

There is also a long and very narrow area in zone 12 that is located along the western border of zone 19, which is closed for fishing for one month after the start of the fishing season. This area is included in zone 12 polygon as well as the southern Gulf polygon.

FORECASTING

A Bayesian serial linear regression model (Appendix 1) was developed to predict recruitment to the snow crab fishery in the southern Gulf of Saint Lawrence (Surette and Wade, 2006). The projections are based on estimated abundances of recruitment stages.

Decisions on future landings are typically based on establishing the quota such that the probability of meeting reference levels are respected. Effort is made to include various sources of uncertainty, including missing data, diffuse priors, and observation errors. Results shows a wide range in fishery recruitment for 2013-2015 with respect to present levels. Cumulative posterior probability plots show the interchange between remaining abundances as well as projected risks under various assumed exploitation rates. All results are presented with associated errors.

DATA

The data used are population and biomass abundance based on kriging estimates for 1997 to 2012. The demographic categories considered are the future recruitment to the fishery, commercial crab remaining after the fishery, and commercial crab landed. Four recruitment stages of male snow crab, which recruit to the fishery in one, two, three and four years are considered. The size limits for these categories were set using a growth model for adolescent male snow crab described by Hebert et al. (2002) as well as other relevant information such as morphometric maturity (Comeau et al. 1998), and minimum legal size. The units for the recruitment category immediately prior to entering the fishery are described in metric tons while all other recruitment categories are in numbers of individuals (Table 4). The commercial crab remaining after the fishery is that part of the exploitable resource which was not caught during the fishery and survived to the survey period the following year. These are legal-sized (carapace width greater than 95 mm), morphometrically mature male snow crab with carapace condition 3, 4 or 5. This last criterion is a qualitative measure of relative age of the carapace. Landed commercial crab abundances are calculated from reported landings (Table 5). Complete nomenclature, classification criteria and descriptions for analytical variables are given in Table 3. Abundance estimates for recruitment were transformed to the logarithmic scale prior to the analysis given that they are assumed to be strictly positive and that their associated error increases with the mean.

MODEL

We relate successive recruitment classes R_{i+1}^{j} and R_{i}^{j-1} using the log-normal linear model which is described in greater detail by (Surette and Wade 2006) as:

$$lnR_{i+1}^{j} = \alpha_j * lnR_i^{j-1} + \beta_j + \epsilon_j \qquad (6)$$

where α_j and β_j are regression coefficients and

$\epsilon \sim N(0, \sigma_j^2)$, where σ_j is the process error.

This model links the recruitment classes in a sequence of log-normal linear regressions: R_{i+1}^{1} is conditioned on R_i^2, R_i^2 is conditioned on R_{i-1}^3 and R_{i-1}^3 is conditioned on R_{i-2}^4. To complete the structure, we specify that R_i^4. be conditioned on itself as part of a first-order autoregressive AR(1) model on the log-scale, where $lnR_{i+1}^4 = \alpha_4 * lnR_i^4 + \beta_4 + \epsilon_4$ and α_4 and β_4 are constants and $\epsilon_4 \sim N(0, \sigma_4^2)$.

The remaining abundance refers to the portion of exploitable part of the total abundance which remains after the fishing season. This portion, added to the one-year recruitment to the fishery

R^1 , provides an index of the total exploitable abundance for the succeeding year and quotas are generally set as a fraction thereof. The goal is to build a predictive model of the remaining abundance for year *i+1* based on the remaining abundance of year *i*, the recruitment to the fishery in year *i* and catches in year *i+1*.

Let Rem_i be the remaining abundance of mature male snow crab with shell condition 3, 4 or 5 for year *i* and let L_{i+1} be biomass caught in year *i+1*. We relate Rem_{i+1} to Rem_i, R_i^1 and L_{i+1} using the linear model:

$$Rem_{i+1} = S * \left(R_i^1 + Rem_i\right) - L_{i+1} \tag{7}$$

where S is a linear regression coefficient.

This model seeks to account for the discrepancy between the predicted total fishable biomass (composed of fisheries recruitment and remaining abundance) and the catches and remaining abundance for the following year. Only the mean of S for the last five years is used for future projections.

The method of estimating the posterior predicted remaining biomass and total biomass for y = 2013 to 2016 as follows:

$$Rem_{2013} = S * (Biomass_{2012}) - L_{2013} \tag{8}$$

$$Biomass_{2013} = R_{2013}^1 + Rem_{2013} \tag{9}$$

$$Rem_{2014} = S * \left(R_{2013}^1 + Rem_{2013}\right) - L_{2014} \tag{10}$$

$$Biomass_{2014} = R_{2014}^1 + Rem_{2014} \tag{11}$$

$$Rem_{2015} = S * \left(R_{2014}^1 + Rem_{2014}\right) - L_{2015} \tag{12}$$

Under this assumed model, one may assess the risk associated with various future quota levels, which is of particular interest to stakeholders and fisheries managers. More precisely, we may calculate the probability that a target exploitation rate $\mathrm{ER}_{\mathrm{target}}$ would be exceeded for hypothetical quota levels for each year from 2013 to 2016. these results are presented graphically in Figure 17 for 2013. It is assumed that once the decision of establishing the maximum risk of exceeding the target exploitation rate is set in 2013, this will establish the expected landings in 2013 which will in turn be used in the predictions and risk analysis for 2014 and so forth.

The probability of exceeding the limit reference level, B_{lim} in year *y,* is obtained by calculating the proportion of the estimates of Rem_y (from equations 7, 8, 10, 12) which are less than or equal to B_{lim}. Similarly, the probability of exceeding an upper stock reference point (B_{usr}) at year *y* is obtained by comparing it to the expected total biomass for the upcoming fishery year (equations 9, 11).

RESULTS AND DISCUSSION

Monte Carlo Markov Chain (MCMC) samples generally showed good mixing (low autocorrelation between samples) and did not exhibit non-stationary behaviour, meaning that no trends were apparent within the ordered samples as they were generated.

Posterior MCMC sample statistics for recruitment to the fishery from 2013 to 2015 are presented in Table 6. The complete data series are shown for each recruitment stage (1 through 4) in Figure 16. Following the observed lower levels of abundance of stages R^4 and R^3, the commercial recruitment predictions for 2013 to 2015 do not show marked increases or decreases. However, the confidence intervals of these estimates are large.

The recruitment model is structured in a similar manner as a population model, with each recruitment stage expressed as a function (linear function on the log-scale) of the preceding stage from the year before. However, the model parameters must not be interpreted as having a simple meaning directly related to specific population dynamic processes. In fact, the identification of recruitment stages from survey data is complicated by the stochastic nature of the growth process, which does not allow us to classify individuals by simple size intervals as we have done here. Our approach is therefore an approximation. Furthermore, there are other processes such as individuals who forego their moult (called skip-moulters) and temporal variability in moult to maturity probabilities, as well as reduced catchability within the sampling gear as individual size decreases, all of which further blurs the definition and interactions between successive recruitment groups. We avoided these complications by focusing rather on the prediction of recruitment by a linear model, which is understood to be an approximation of a very complex process and thus avoided making specific interpretations of its parameters as they relate to specific processes within the population.

REFERENCES

Armstrong, M., Renard, D., Rivoirard, J., and Petitgas, P. 1992. Geostatistics for fish survey data. Course publicized by ICES, Centre de Geostatistique, Fontainebleau.

Comeau, M., Conan, G.Y., Maynou, F., Robichaud, G., Therriault, J.-C., and Starr, M. 1998. Growth, spatial distribution and abundance of benthic stages of the snow crab (*Chionoecetes opilio*) in Bonne Bay, Newfoundland, Canada. Can. J. Fish. Aquat. Sci. 55: 262-279.

Cressie, N. 1993. Statistics for Spatial Data (Wiley Series in Probability and Statistics). Revised edition. Wiley-Interscience.

Delhomme, J.P. 1978. Kriging in the hydrosciences. Advances in Water Resources 1: 251-266.

Deutsch, C.V., and Journel, A.G. 1992. GSLIB: Geostatistical software library and users guide. New York: Oxford University Press. Oxford.

DFO. 2006. Proceedings of the Assessment Framework Workshop on Southern Gulf of St. Lawrence Snow Crab (Areas 12, E, F and 19), Gulf Regional Advisory Process; 11-14 October 2005. DFO. Can. Sci. Advis. Sec. Proceed. Ser. 2006/042.

DFO. 2012. Proceedings of the Gulf Region Science Peer Review Framework Meeting of Assessment Methods for the Snow Crab Stock of the southern Gulf of St. Lawrence. DFO Can. Sci. Advis. Sec. Proc. Ser. 2011/023. 50 p.

Dutil, J.D. 2011. A hierarchical classification of the seabed based on physiographic and oceanographic features in the St. Lawrence. Can. Tech. Rep. Fish. Aquat. Sci.##.

Efron, B. 1979. Bootstrapping methods: Another look at the jackknife. Ann. Stat. 7: 1-26.

Goovaerts, P. 1997. Geostatistics for natural resources evaluation. Applied geostatistics series. Oxford University Press.

Hebert, M., Benhalima, K., Miron, G., & Moriyasu, M. 2002. Moulting and growth of male snow crab, *Chionoecetes opilio* (O.Fabricius,1788) (Decapoda,Majidae) In the southern Gulf of St. Lawrence. Crustaceana 75: 671—702.

Hengl, T., Heuvelink, G.B.M., and Rossiter, D.G. 2007. About regression-kriging, from equations to case studies. Computers and Geosciences 33: 1301-1315.

Isaaks, E.H., and Srivastava, R.M. 1989. An Introduction to Applied Geostatistics. Vol. 46. Oxford University Press.

Keitt, T.H., Bivand, R., Pebesma, E., and Rowlingson, B. 2012. rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.7-8.

Landry, J.-F., Wade, E., Moriyasu, M., and Hébert, M. 2013. Summary of activities during the 2012 snow crab trawl survey in the southern Gulf of St. Lawrence. DFO Can. Sci. Advis. Sec. Res. Doc. 2013/115.

Matheron, G. 1963a. Principles of geostatistics. Economic geology 58: 1246-1256.

Matheron, G. 1963b. Traité de Géostatistique appliqué. Editions Technique v. 2., Paris School of Mines. Cah. Cent. Morphol. Math., Fontainbleau.

Moriyasu, M., Wade, E., Hébert, M., and Biron, M. 2008. Review of the survey and analytical protocols used for estimating abundance indices of southern Gulf of St. Lawrence snow crab from 1988 to 2006. DFO Can. Sci. Advis. Sec. Res. Doc. 2008/069.

Pebesma, E.J. 2004. Multivariable geostatistics in S: the gstat package. Computers and Geosciences 30: 683-691.

R Core Team. 2012. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

Rivoirard, J., Simmonds, J., Foote, K.G., Fernandes, P., and Bez, N. 2000. Geostatistics for Estimating Fish Abundance, Wiley, Blackwell Science

Rossi, R., Mulla, D., Journel, A., and Franz, E. 1992. Geostatistical Tools for Modeling and Interpreting Ecological Spatial Dependence. Ecological Monographs 62: 277-314.

Surette, T.J., and Wade, E. 2006. Bayesian serial linear regression models for forecasting the short-term abundance of commercial snow crab *Chionoecetes opilio*. Can. Tech. Rep. Fish. Aquat. Sci. 2672.

Surette, T.J., and Wade, E. 2012. New sampling design for the 2012 southern Gulf of Saint Lawrence snow crab (*Chionoecetes opilio)* survey. Can. Tech. Rep. Fish. Aquat. Sci. (In prep.).

Surette, T.J., Marcotte, D., and Wade, E. 2007. Predicting snow crab (*Chionoecetes opilio*) abundance using kriging with external drift with depth as a covariate. Can. Tech. Rep. Fish. Aquat. Sci. 2763.

Webster, R., and Olivier, M.A. 2001. Geostatistics for environmental scientists. John Wiley and Sons, New York.

TABLES

Table 1: Estimated commercial biomass, recruitment biomass, and residual biomass (t, mean and 95% confidence interval range) of commercial-sized adult male snow crab, Chionoecetes opilio, in the southern Gulf of St. Lawrence, 1997 to 2012. Recruitment refers to snow crab with carapace conditions 1 and 2 whereas residual biomass refers to snow crab with carapace conditions 3 to 5.

Survey year	Biomass	Recruitment	Residual biomass
1997	65310	37619	27690
	(54801-77239)	(26376-52064)	(21995-34407)
1998	57595	29818	27775
	(45630-71735)	(17580-47435)	(21022-36013)
1999	57051	25874	31177
	(47946-67376)	(15918-39818)	(25051-38346)
2000	49823	39845	9977
	(40473-60682)	(30543-51093)	(6649-14401)
2001	59150	42243	16905
	(47740-72460)	(31198-55942)	(12657-22125)
2002	79559	66481	13075
	(66688-94181)	(53434-81746)	(10451-16157)
2003	84423	57503	26919
	(71964-98410)	(44809-72679)	(21223-33674)
2004	103429	83702	19726
	(91029-117036)	(70955-98069)	(15836-24280)
2005	82537	58398	24140
	(73487-92387)	(48417-69824)	(18726-30632)
2006	74285	54371	19914
	(66192-83087)	(46124-63660)	(16161-24275)
2007	66660	39635	27025
	(60183-73638)	(33089-47092)	(23354-31106)
2008	52564	31555	21010
	(46658-59006)	(25181-39048)	(17960-24426)
2009	30920	20520	10399
	(27237-34959)	(16848-24754)	(8560-12516)
2010	35795	20351	15444
	(31681-40291)	(15360-26450)	(12859-18394)
2011	63162	29394	33768
	(55965-71022)	(20909-40190)	(28297-39985)
2012	74997	48969	26028
	(65822-85086)	(38667-61173)	(21950-30641)

Table 2: Commercial biomass estimates (t, mean and 95% confidence interval) by snow crab fishing area 12, 19, 12E, and 12F.

Year	Area12	Area19	Area 12E	Area 12F
1997	57617 (48009 - 68577)	4779 (3521 - 6342)	1283 (391 - 3168)	1260 (362 - 3217)
1998	48428 (37573 - 61444)	6073 (4581 - 7897)	1096 (226 - 3319)	1492 (430 - 3798)
1999	46703 (38433 - 56221)	6113 (4924 - 7502)	1742 (763 - 3433)	1778 (770 - 3528)
2000	39021 (31273 - 48103)	7040 (6181 - 7984)	626 (141 - 1818)	2645 (1870 - 3634)
2001	49273 (39755 - 60379)	4773 (3780 - 5947)	1111 (416 - 2422)	3465 (2576 - 4563)
2002	68367 (57680 - 80451)	5465 (4351 - 6779)	1150 (403 - 2613)	3327 (2368 - 4547)
2003	71259 (60949 - 82804)	8275 (7164 - 9509)	1139 (405 - 2566)	2690 (1784 - 3898)
2004	95281 (84929 - 106538)	4486 (3380 - 5841)	1017 (297 - 2574)	1853 (952 - 3267)
2005	75050 (66711 - 84136)	3939 (2796 - 5395)	958 (244 - 2612)	1681 (713 - 3385)
2006	67788 (60571 - 75622)	4191 (2775 - 6081)	805 (143 - 2614)	1022 (263 - 2767)
2007	58954 (53188 - 65169)	5434 (4232 - 6872)	544 (78 - 1940)	1014 (342 - 2359)
2008	47922 (42681 - 53625)	3210 (2189 - 4547)	293 (15 - 1459)	642 (145 - 1862)
2009	25895 (22669 - 29449)	3435 (2872 - 4076)	271 (39 - 964)	1088 (682 - 1650)
2010	29366 (25778 - 33309)	4953 (4329 - 5641)	281 (35 - 1059)	726 (331 - 1393)
2011	51381 (45110 - 58274)	8346 (7245 - 9565)	705 (162 - 2023)	1900 (1135 - 2993)
2012	64238 (56254 - 73031)	7668 (5944 - 9736)	577 (68 - 2214)	1450 (480 - 3409)

Table 3: Variable names, descriptions, units, and classification criteria for snow crab life stages. "cw" is carapace width.

Variable	Description	Units	Classification Criteria
R^1	Recruitment abundance to the fishery in one year	t x1000	Morphometrically mature males, cw >= 95 mm, shell condition 1 and 2 (recently terminal moulted)
R^2	Recruitment abundance to the fishery in two years	number x 1,000,000	Morphometrically non-mature males, cw >= 83mm
R^3	Recruitment abundance to the fishery in three years	number x 1,000,000	Morphometrically non-mature males, 69 mm <= cw < 83 mm
R^4	Recruitment abundance to the fishery in four years	number x 1,000,000	Morphometrically non-mature males, 56 mm <= cw < 69 mm
Rem	Remaining fishable abundance after the fishery	t x1,000	Morphometrically mature males, cw >= 95 mm, shell condition 3,4 and 5 (older carapace)
L	Landings	t x1,000	

Table 4: Mean abundance or biomass (standard deviation) by recruitment stage obtained from kriging for the survey years, 1997 to 2012. The number in superscript refers to the number of years to recruitment to the fishery.

Survey year	R^4 (number, million)	R^3 (number, million)	R^2 (number, million)	R^1 (biomass, 1000 t)
1997	114.17 (18.68)	92.72 (12.22)	57.87 (7.05)	37.61 (6.57)
1998	139.54 (19.96)	91.57 (11.52)	57.12 (9.31)	29.81 (7.67)
1999	199.71 (26.21)	150.90 (16.17)	115.03 (19.01)	25.87 (6.13)
2000	238.70 (21.44)	159.44 (14.36)	89.33 (8.40)	39.84 (5.25)
2001	313.24 (22.27)	229.18 (15.24)	135.69 (14.68)	42.24 (6.32)
2002	166.74 (16.16)	241.85 (18.07)	199.68 (16.89)	66.48 (7.23)
2003	137.82 (15.87)	207.09 (18.08)	181.43 (13.15)	57.50 (7.12)
2004	86.37 (7.86)	122.76 (9.30)	142.52 (10.01)	83.70 (6.92)
2005	63.27 (5.87)	79.40 (6.02)	117.09 (10.25)	58.39 (5.46)
2006	54.96 (5.73)	49.78 (3.40)	65.31 (8.00)	54.37 (4.47)
2007	57.24 (6.30)	47.87 (4.90)	55.98 (6.11)	39.63 (3.57)
2008	80.36 (6.33)	54.32 (4.55)	45.83 (6.00)	31.55 (3.54)
2009	89.41 (5.51)	69.45 (5.44)	43.60 (4.72)	20.52 (2.01)
2010	140.42 (7.75)	109.14 (8.73)	71.77 (6.23)	20.35 (2.83)
2011	91.48 (7.40)	98.67 (7.18)	87.60 (7.49)	29.39 (4.93)
2012	96.04 (8.64)	86.84 (9.81)	80.52 (7.97)	48.96 (5.75)

Table 5: Mean biomass (standard deviation) of post-fishery remaining abundance (residual biomass) obtained from kriging and annual fishery landings for 1997 to 2012.

Survey Year	Residual biomass (1000 t)	Landings (1000 t)
1997	27.69 (3.17)	17.65
1998	27.77 (3.83)	13.86
1999	31.17 (3.39)	15.51
2000	9.97 (1.98)	19.18
2001	16.90 (2.42)	18.51
2002	13.07 (1.45)	26.17
2003	26.91 (3.18)	21.16
2004	19.72 (2.15)	31.66
2005	24.14 (3.04)	36.07
2006	19.91 (2.07)	29.12
2007	27.02 (1.97)	27.16
2008	21.01 (1.65)	24.89
2009	10.39 (1.01)	23.99
2010	15.44 (1.41)	9.549
2011	33.76 (2.98)	10.67
2012	26.02 (2.21)	21.96

Table 6: Posterior MCMC sample statistics (mean, standard deviation, 2.5th percentile, median, 97.5th percentile, and MC error) for recruitment biomass (R1; X 1000 t) predictions for 2013 to 2015.

Year	Mean	Std. Dev.	2.5th percentile	median	97.5th percentile	MC error
2013	40.25	4.645	31.55	40.02	50.27	0.09233
2014	35.58	5.155	26.2	35.29	46.76	0.07098
2015	36.27	6.31	25.2	35.78	50.37	0.073

FIGURES

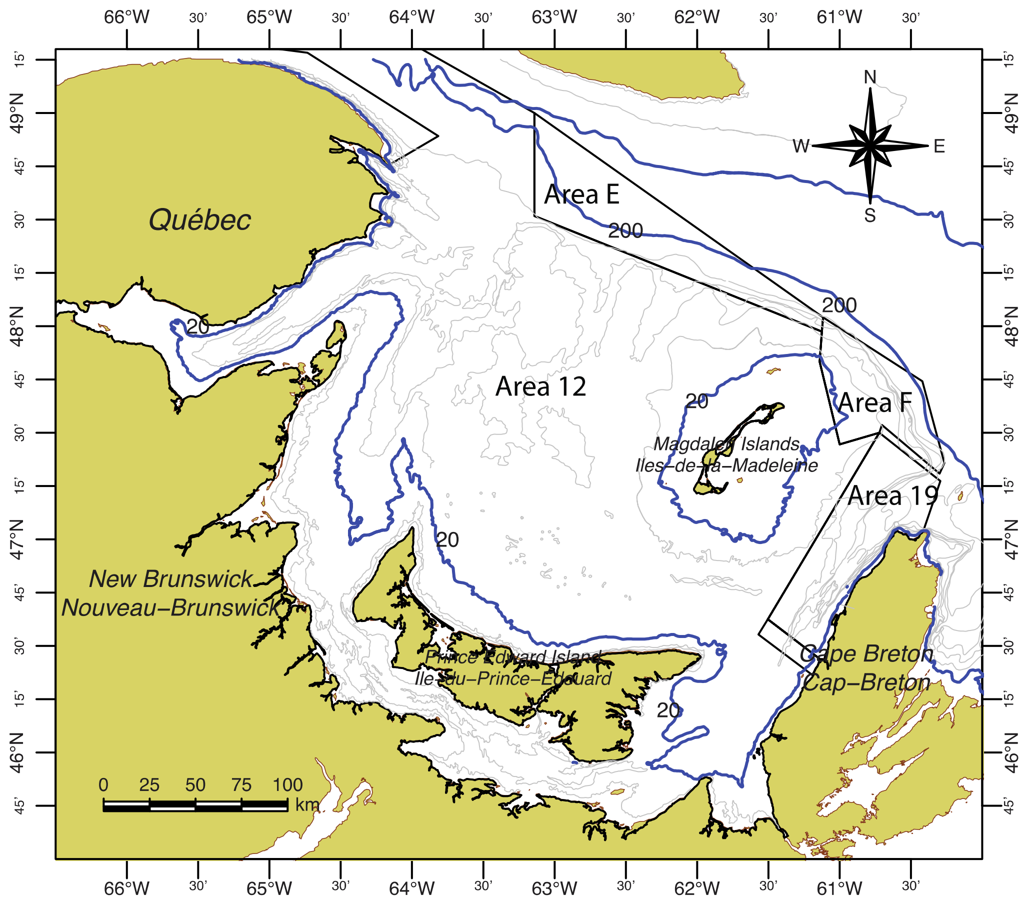


Figure 1: Locations of fishing grounds and snow crab fishery zones in the southern Gulf of St. Lawrence.

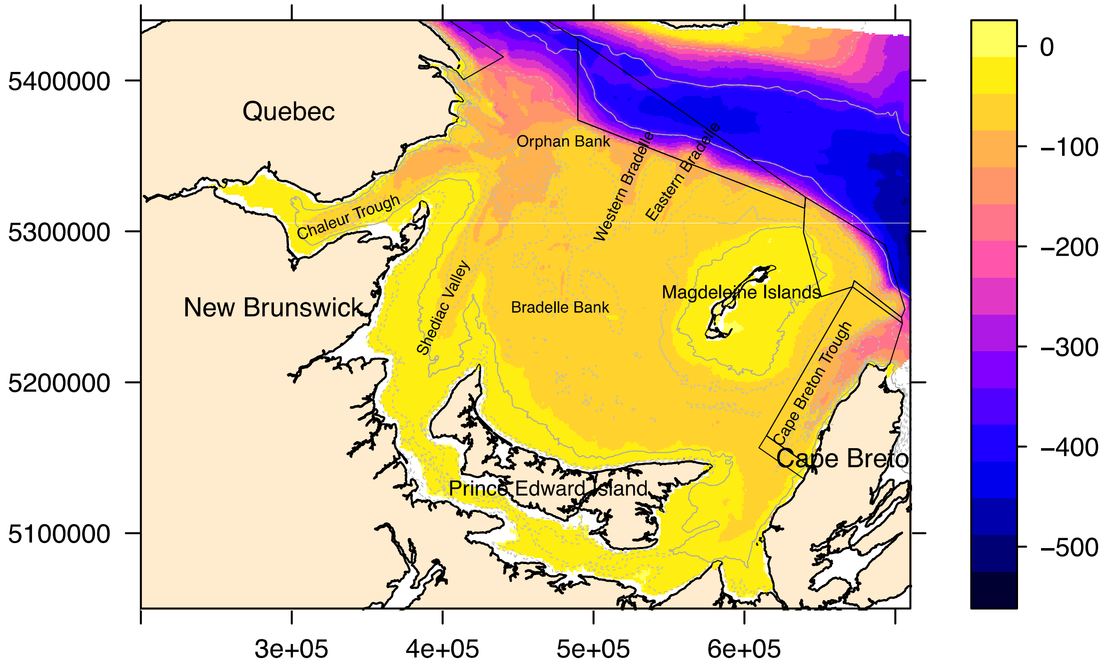


Figure 2: Depth in the southern Gulf of St. Lawrence. Depth units are in meters and coordinates are in meters using the Universal Transverse Mercater coordinate system (UTM) and the NAD 83 (Zone 20) datum.

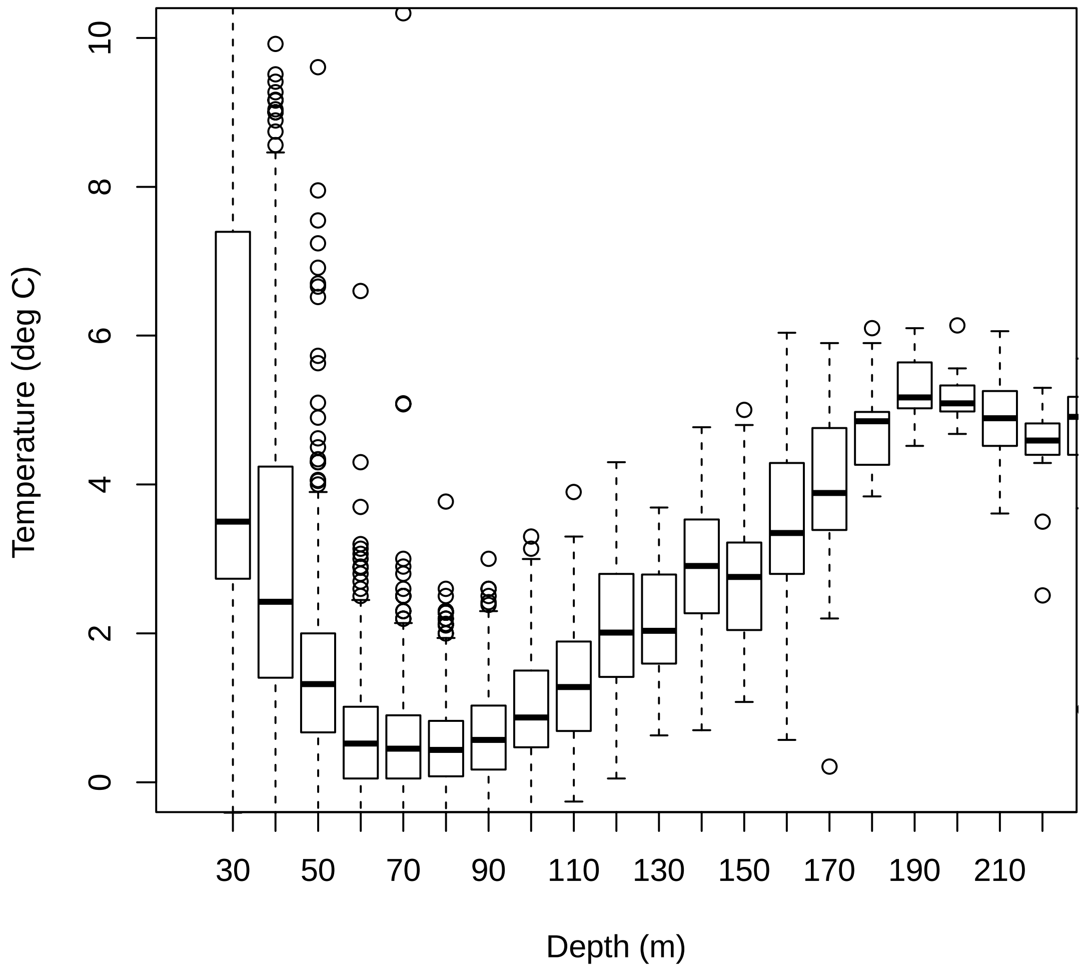


Figure 3: Boxplots of the measured temperature (ºC) by depth bins (m) in the southern Gulf of St. Lawrence, 1998 to 2012.

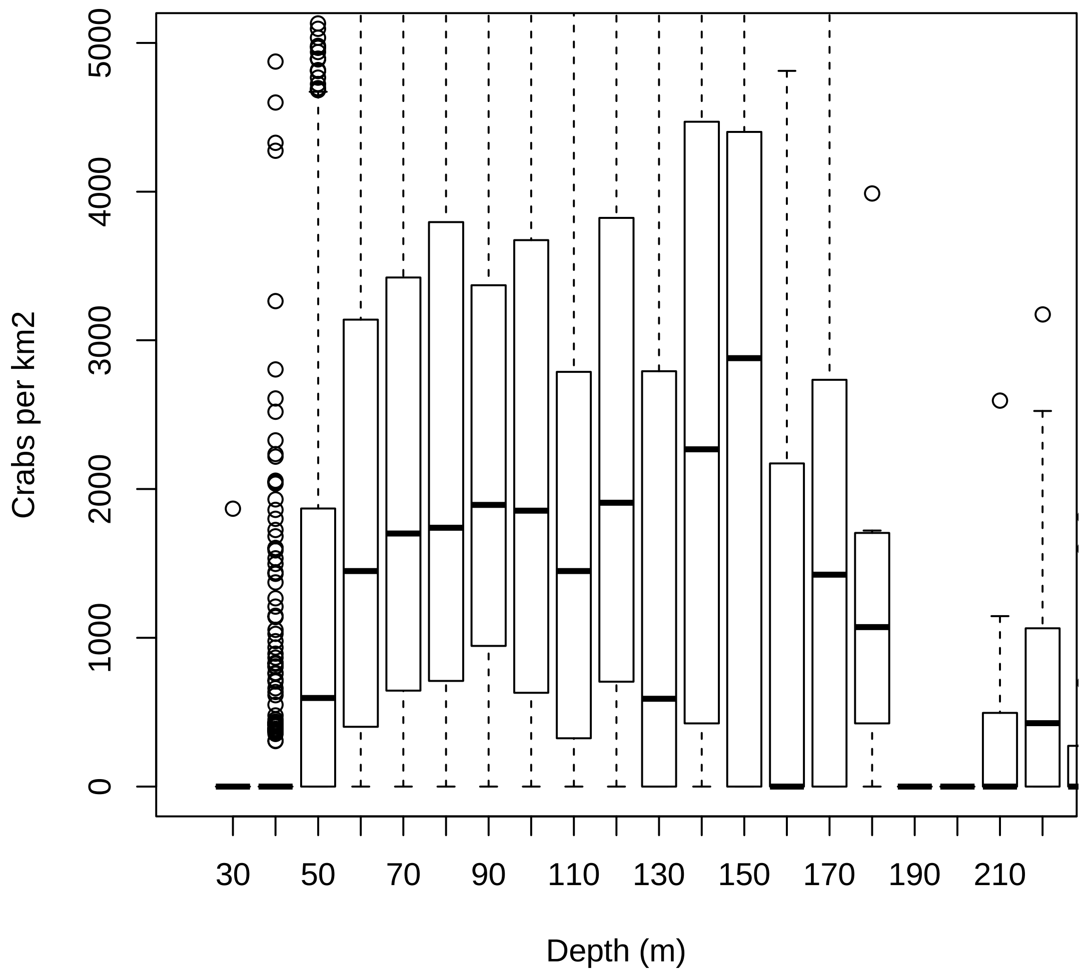


Figure 4: Boxplots of estimated commercial-size (>= 95 mm carapace width) adult male crab densities (number per km^2)by depth bins (m) in the southern Gulf of St. Lawrence, 1998 to 2012.

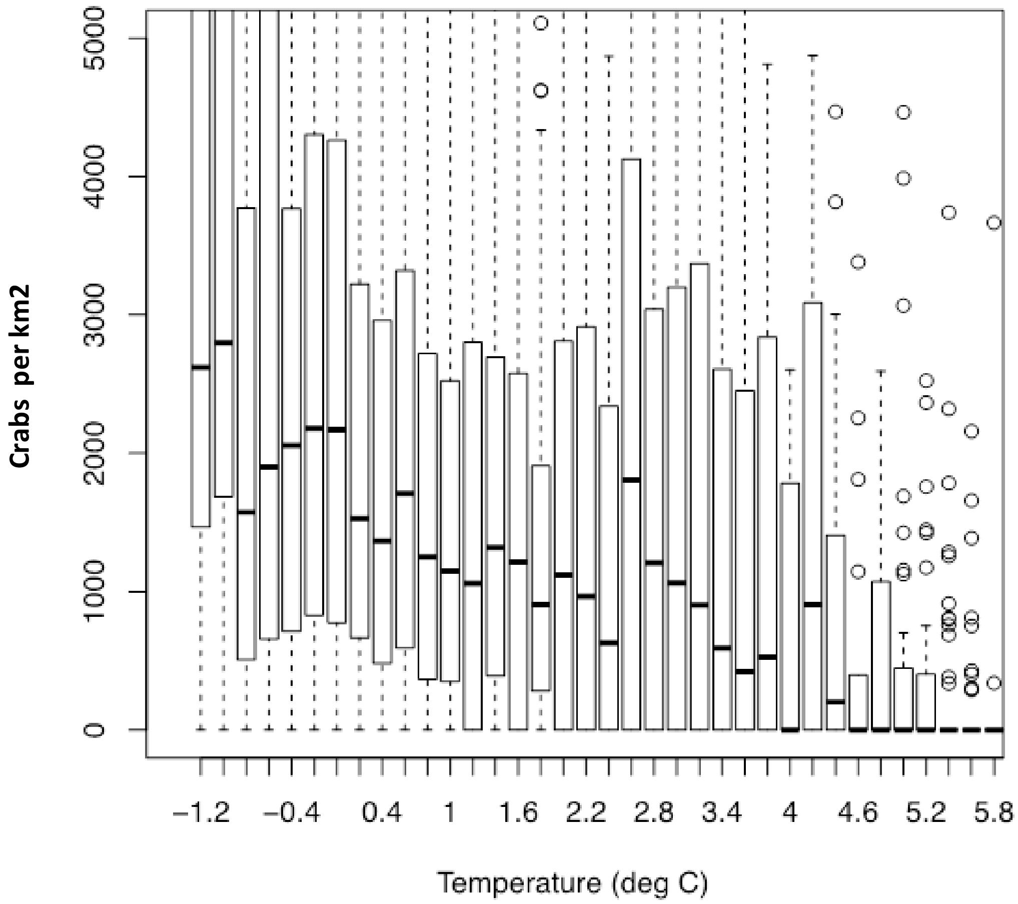


Figure 5: Boxplots of estimated commercial-size (>= 95 mm carapace width) adult male crab densities (number per km2) by temperature bins (ºC) in the southern Gulf of St. Lawrence, 1998 to 2012.

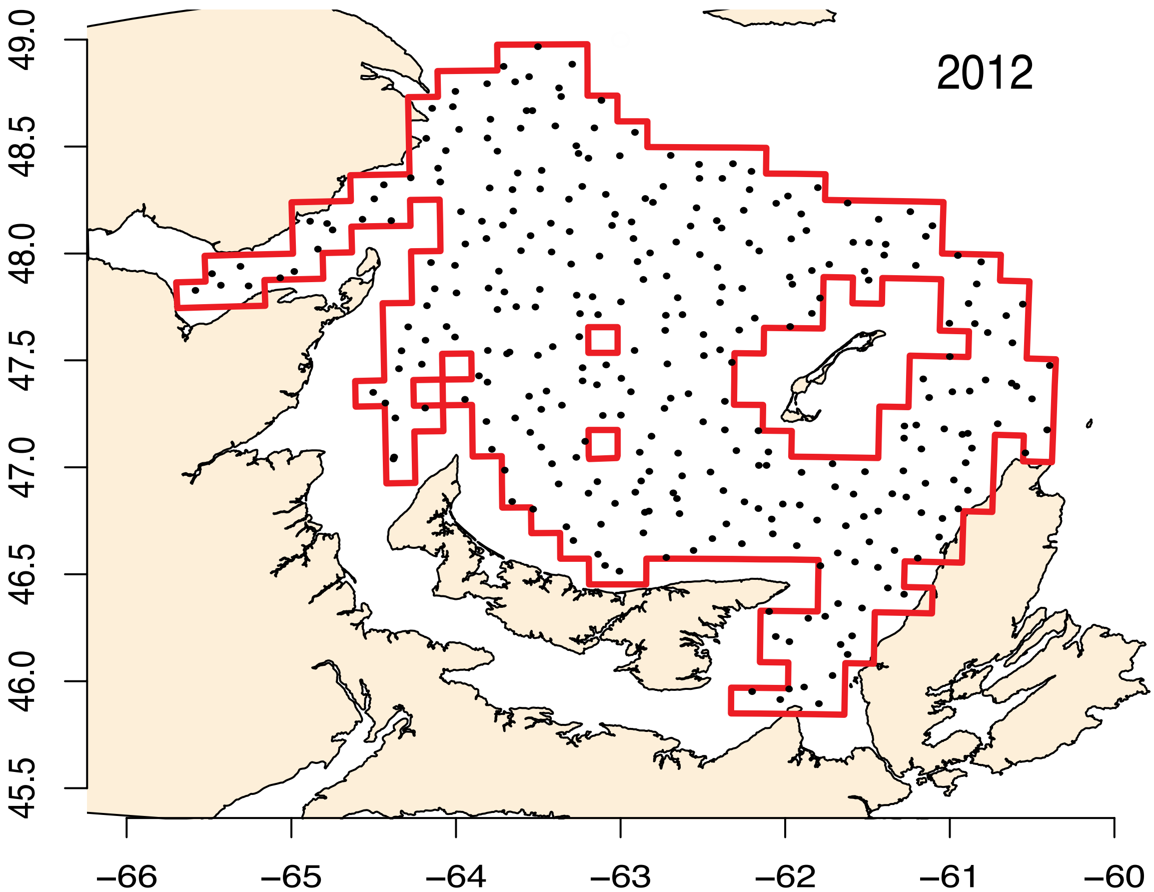


Figure 6: Envelope of sampling coverage for 2012. The empty red squares are grid cells with no sample in 2012.

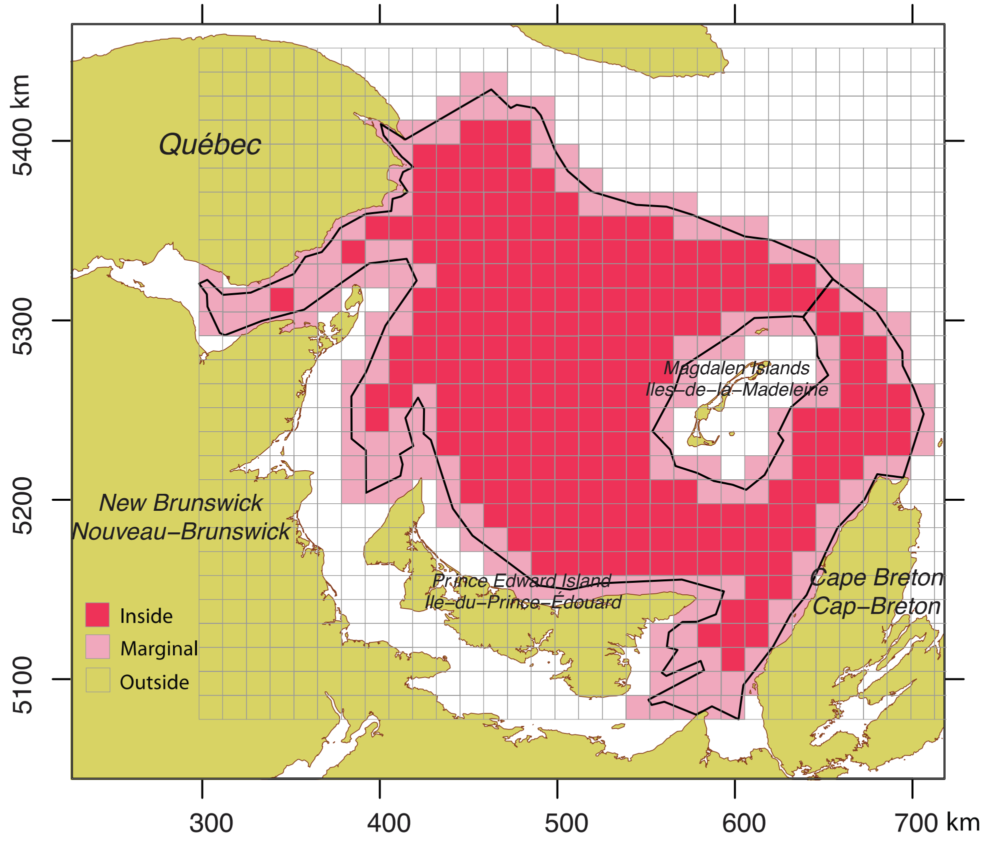


Figure 7: Grid array defined by 13.36 km by 13.36 km squares (n=325) superimposed over the survey area. Shown in red are grid cells entirely within the survey area, in white are those outside the survey area and in pink are grid cells lying on the margins. The projection is in UTM (NAD83, zone 20).

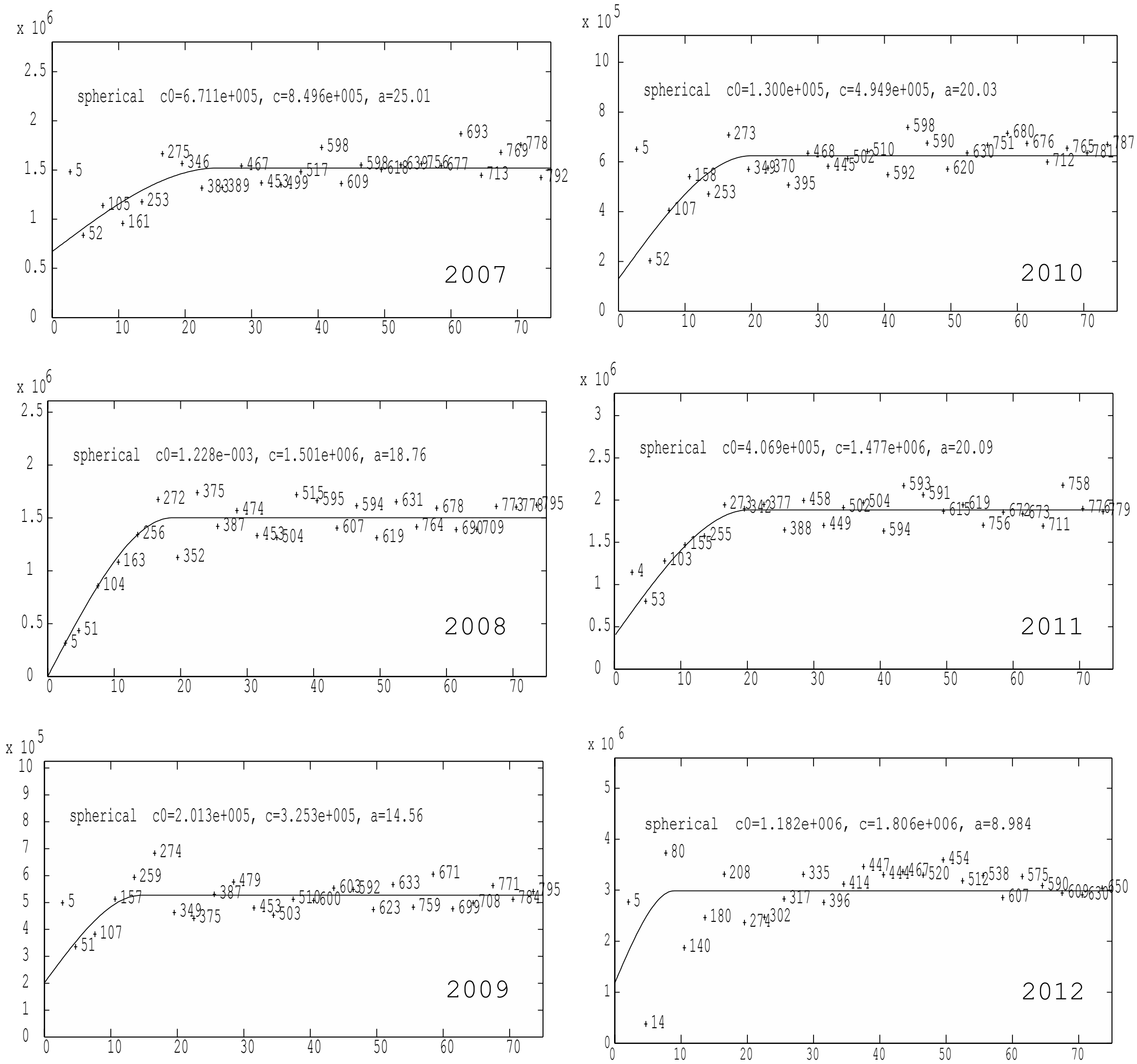


Figure 8: Annual variogram models for the commercial sized adult male crab category from 2007 to 2012.

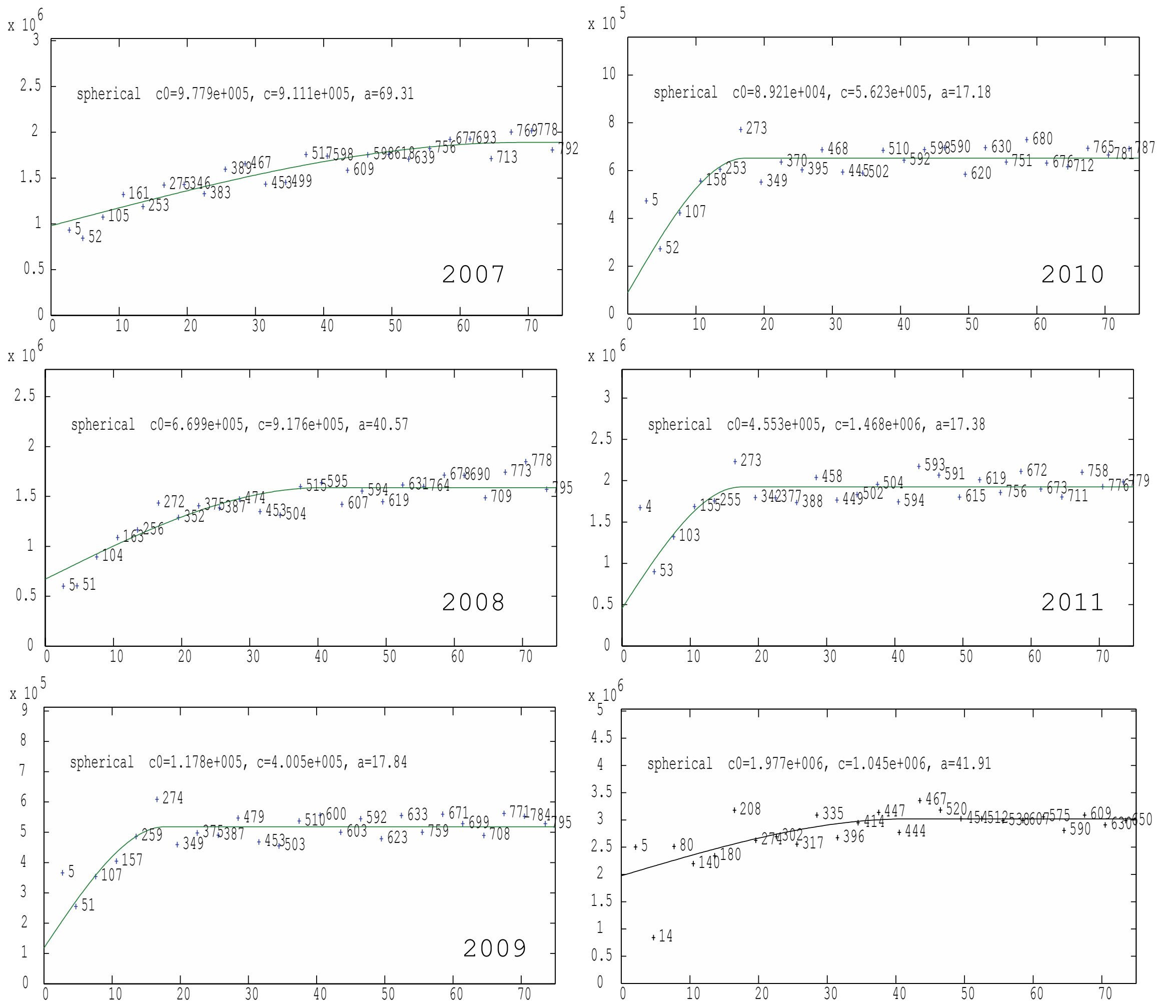


Figure 9: Variogram models averaged over 3 years for the commercial sized adult male crab category from 2007 to 2012.

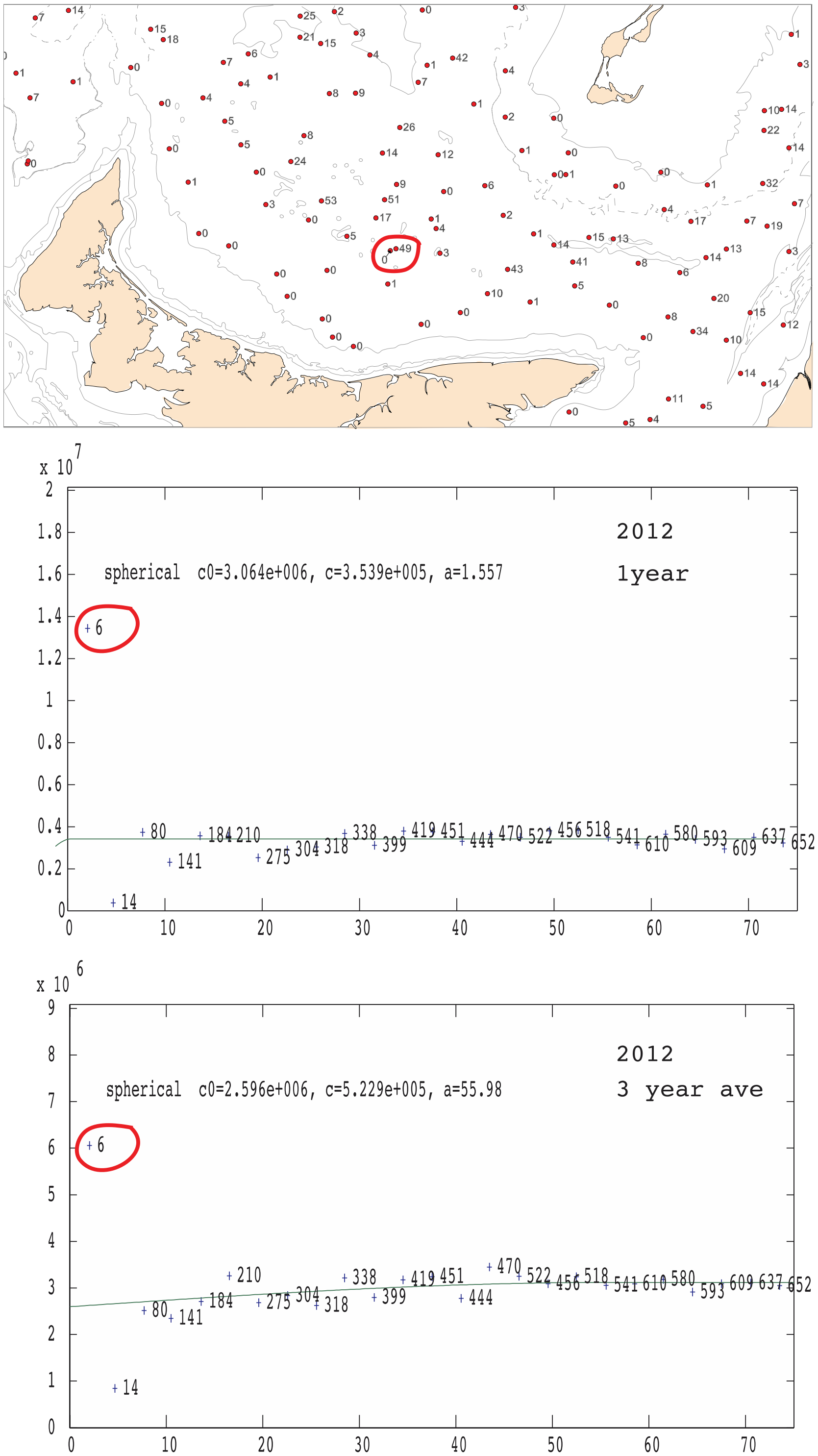


Figure 10: Location of the outlier sample in the 2012 survey (upper panel) and the effects of including the outlier on the annual variogram (middle panel) and the 3-year average variogram (lower panel) for the commercial sized adult male crab category in 2012.

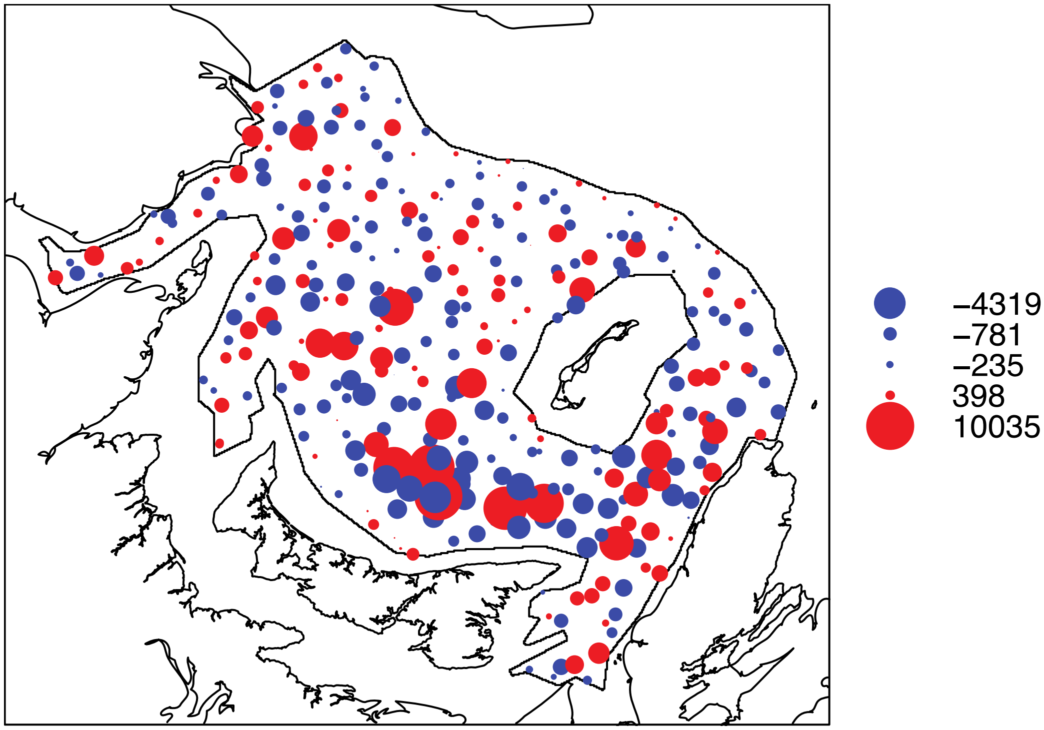


Figure 11: Crossvalidation map for commercial sized adult male crab density estimates in 2012. The density estimates were obtained using KED with depth, a 3-year average variogram, and catches expressed as weights at each station.

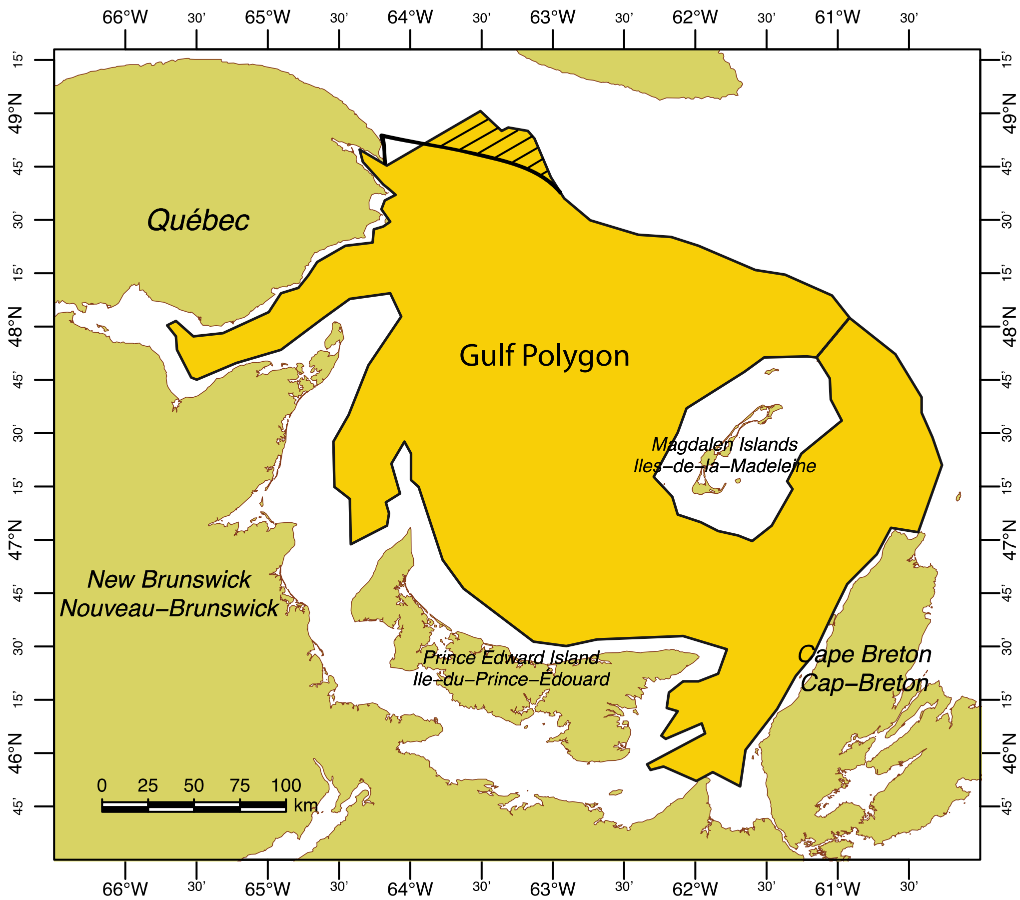


Figure 12: Snow crab sampling and estimation area (shown in yellow) covering the 20 to 200 fathom isobaths representing an area of 57,840 km2. The heavy black line in the northern part of the sampling area shows the limit of the previous survey. The hash pattern shows the newly covered area.

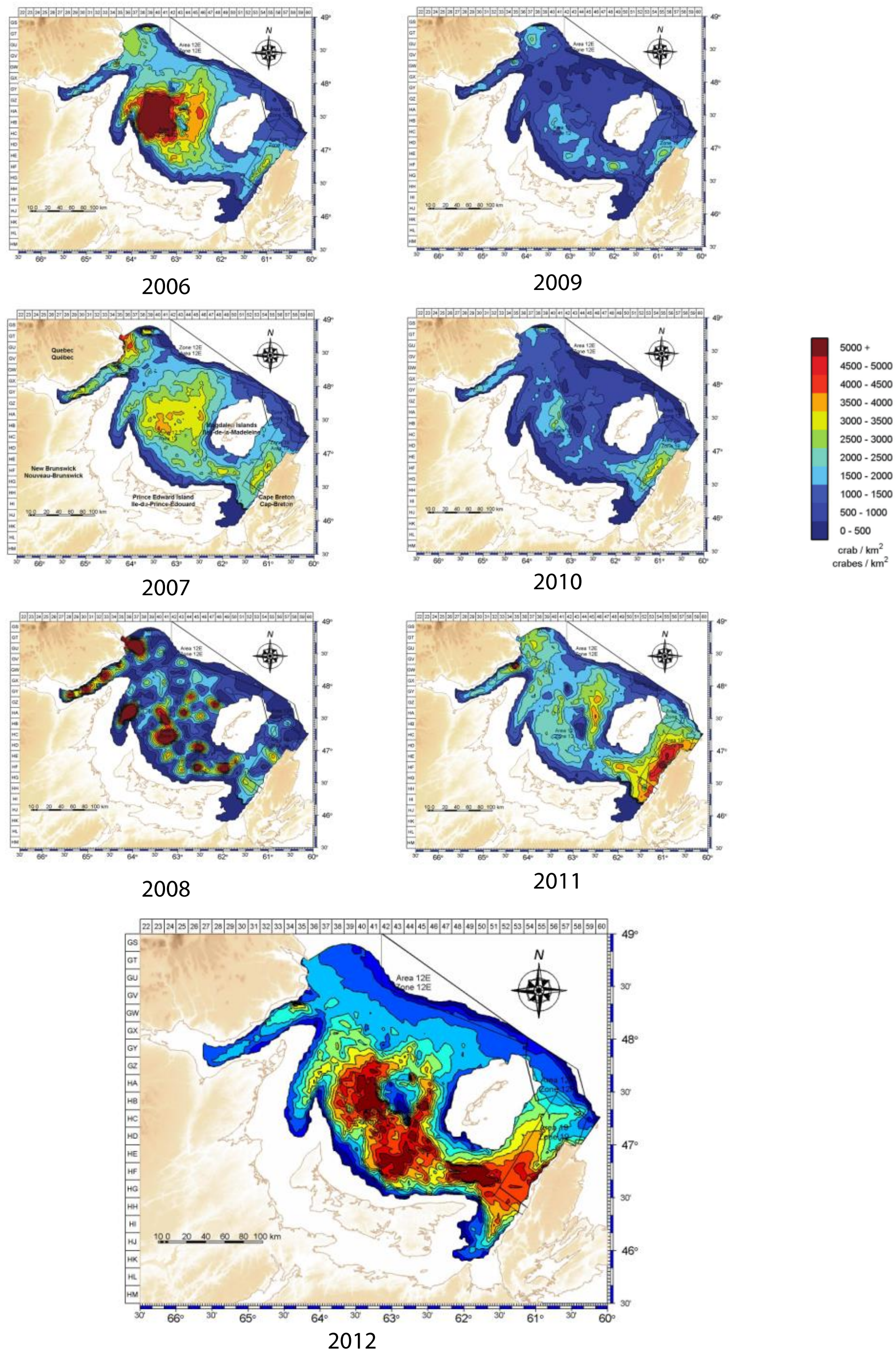


Figure 13: Density (number of crab per km2) contours of commercial-sized (>= 95 mm CW) adult male snow crab based on trawl survey data in the southern Gulf of St. Lawrence, 2006 to 2012.

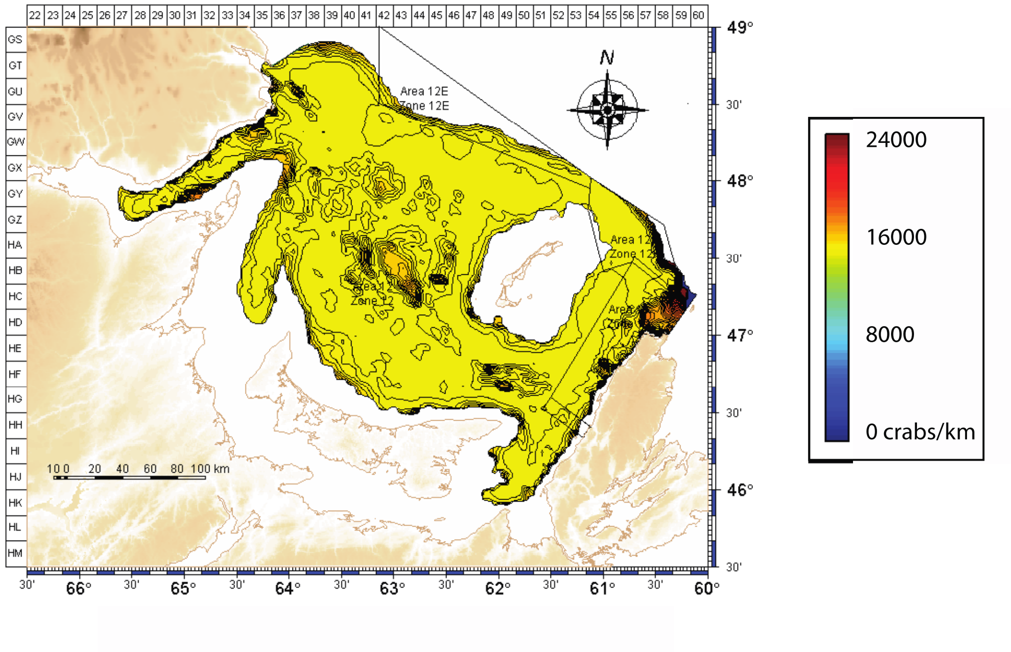


Figure 14: Error map (kriging variance) for the commercial sized adult male crab category in 2012.

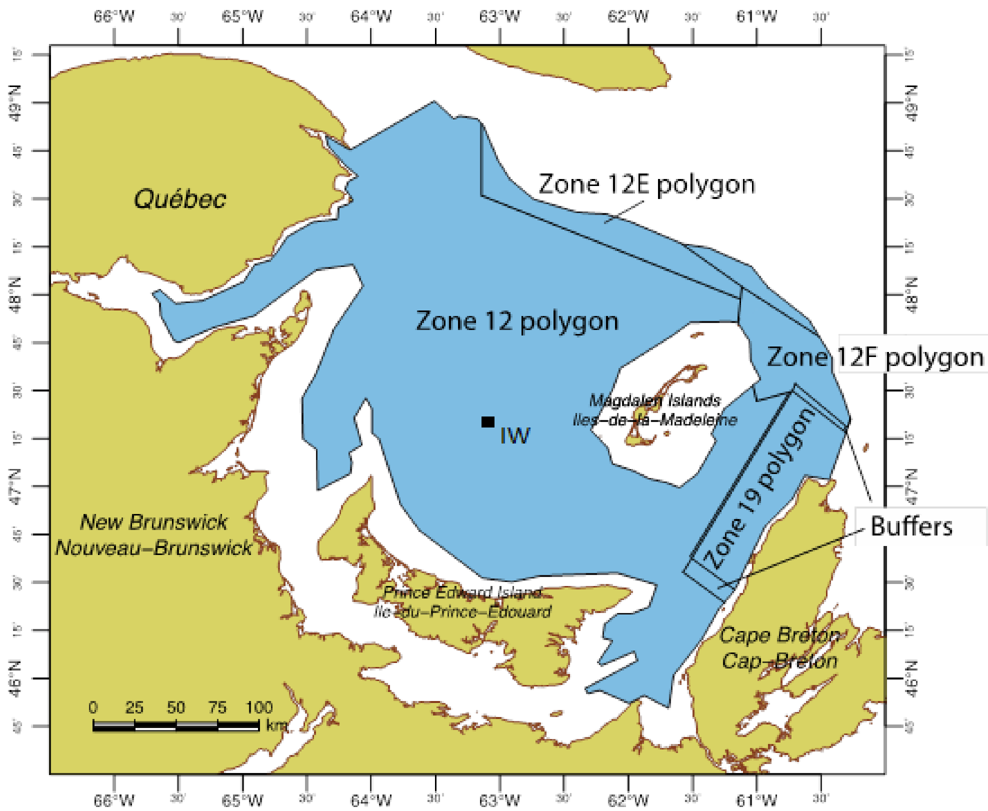


Figure 15: Delimitation of kriging polygons used to estimate the biomass in the snow crab fishing areas 12, 19, 12E, and 12F. Also shown are the buffers which are closed to fishing and the closed area associated with the Irving Whale (IW) site.

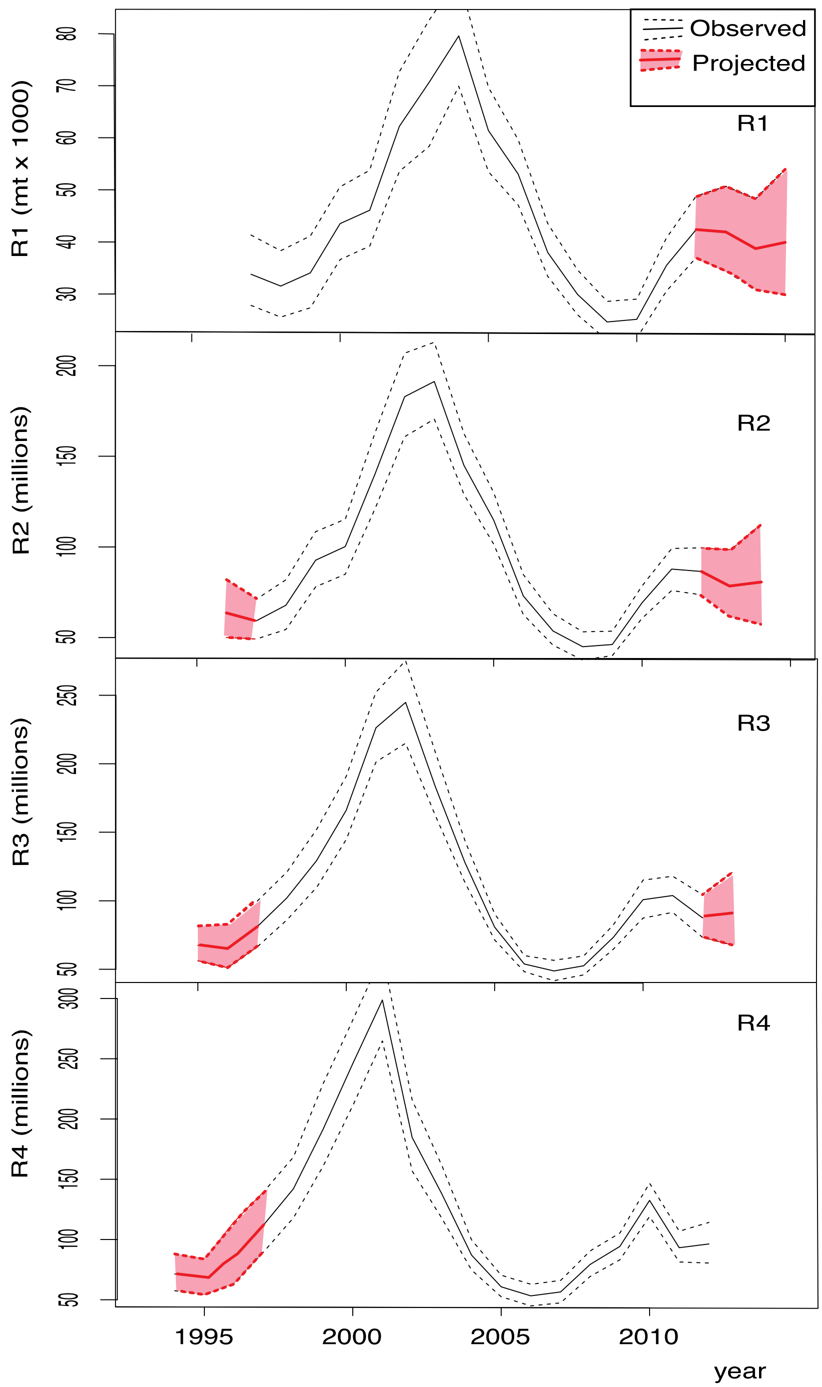


Figure 16: Estimated (without shading; 1997 to 2012) and predicted (red shading; 1994 to 1996, 2013 to 2016) mean and 95% confidence interval range of the R4, R3, R2, and R1 male snow crab stages in the southern Gulf of St. Lawrence. The predictions are based on the Bayesian model.

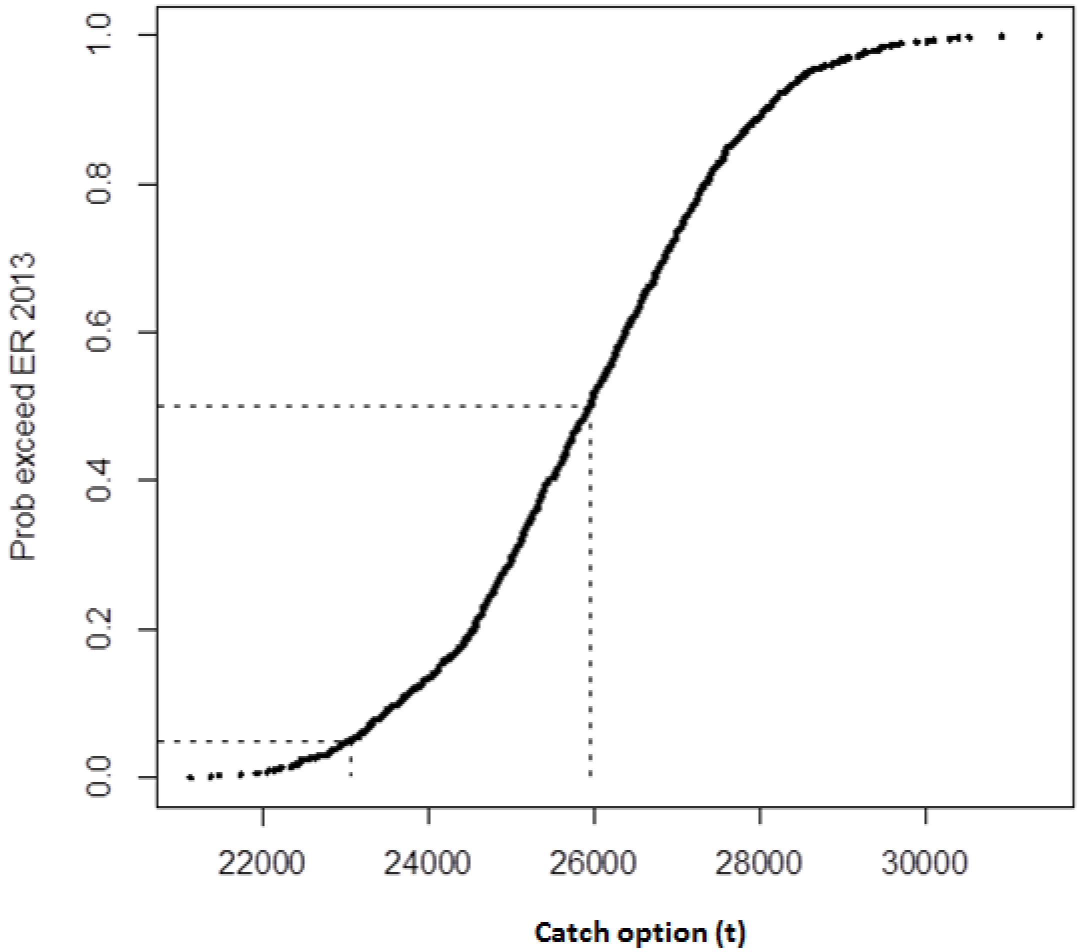


Figure 17: Probability of exceeding the exploitation removal reference of 0.346 in 2013 for different quota options for the 2013 fishery. Dashed lines indicate quota options corresponding to 0.05 and 0.5 probabilities.

APPENDICES

APPENDIX 1. WINBUGS CODE FOR THE BAYESIAN RECRUITMENT MODEL.

```
#  - Non-hierarchical slope parameter only.
#  - Commom AR(1) parameters.
#  - Hierarchical S.
#  - Data from 1997-2012.
# The model predicts k (=4) years before and after the time
# series for most variables. If we focus on prediction
# the indices of the time series elements range from k+1 to
# N + 2*k, where k is the number of recruitment classes and N
# is the number of years in the data time series. For N = 16
# and k = 4, this corresponds to indices 5 through 24. The last
# data year (2012) has an index value of 20 (=N+k). To isolate
# R1's from the start of the data series plus 4 predicted years,
# we would extract R[5:24, 1] (= R[(k+1):(N+2*k), 1]).
# DIC = -67.36
model{
 # Provide process parameters prior distributions:
  for  (j in 1:3){
    alpha[j] ~ dnorm(0, 0.0001)
  }
  for  (j in 1:4){
    # Array of process precision parameters:
    tau.process.r[j] ~ dgamma(0.0001, 0.0001)
    sigma.process.r[j] <- 1 / sqrt(tau.process.r[j])
  }

  # Provide informative priors for missing kriging standard errors:
  for (i in 1:k){
    # dgamma(0.6, 0.1) # Old prior.
    for (j in 1:k){
      sigma.r[i,j] ~ dlnorm(-2.123902, 5.525642)
      sigma.r[i+N+k,j] ~ dlnorm(-2.123902, 5.525642)
    }
  }

  # Calculate kriging precisions from defined standard errors:
  for (i in 1:(N+(2*k))){
    for (j in 1:4){
      tau.r[i,j] <- 1 / pow(sigma.r[i,j],2)
    }
  }

  # Draw log-recruitment priors:
  for (j in 1:k){
    log.R[1,j] ~ dnorm(0, 0.0001)
  }

  # Autoregressive AR(1) prior for R4's:
  phi.alpha ~ dnorm(0, 0.0001)
  phi.beta ~ dnorm(0, 0.0001)
  for(i in 2:(N+(2*k))){
    # Define process with error
    mu.temp.r[i,4] <- phi.alpha * log.R[i-1,4] + phi.beta
    log.R[i,4] ~ dnorm(mu.temp.r[i,4], tau.process.r[4])

    # Observation error
    mu.r[i,4] ~ dnorm(log.R[i,4], tau.r[i,4])
  }
```

```
# State regression models:
for (j in 1:3){
  for(i in 2:(N+(2*k))){
    # Define process with error
    mu.temp.r[i,4-j] <- alpha[4-j] * log.R[i-1,5-j]
    log.R[i,4-j] ~ dnorm(mu.temp.r[i,4-j], tau.process.r[4-j])

    # Observation error
    mu.r[i,4-j] ~ dnorm(log.R[i,4-j], tau.r[i,4-j])
  }
}

# Convert log-abundances to true scale:
for(i in 1:(N+(2*k))){
  for (j in 1:k){
    R[i,j] <- exp(log.R[i,j])
  }
}

# Survivorship coefficient:
mu.S ~ dnorm(0, 0.0001)
tau.S ~ dgamma(0.0001, 0.0001)
for (i in 1:(N+(2*k))){
  S[i] ~ dnorm(mu.S, tau.S)
}

# Remaining process error:
tau.process.rem ~ dgamma(0.0001, 0.0001)

# Calculate precision:
for (i in 1:k){
  sigma.rem[i] ~ dlnorm(1.08963, 2.74476)
  sigma.rem[N+k+i] ~ dlnorm(1.08963, 2.74476)
}
for (i in 1:(N+(2*k))){
  tau.rem[i] <- 1 / pow(sigma.rem[i],2);
}

mu.rem[1] ~ dnorm(0, 0.0001)

# Define initial remaining abundance according to kriged estimate
Rem[k+1] ~ dnorm(mu.rem[k+1], tau.rem[k+1])

# Create independent R variable:
for (i in 1:(N+(2*k))){
  for (j in 1:4){
    cR[i,j] <- cut(R[i,j])
  }
}

# Model remaining biomasses:
for (i in (k+2):(N+k)){
  mu.temp.rem[i] <- S[i] * (cR[i-1,1]+ Rem[i-1]) - L[i]
  Rem[i] ~ dnorm(mu.temp.rem[i], tau.process.rem)

  # Observation error
  mu.rem[i] ~ dnorm(Rem[i], tau.rem[i])
}

# Calculate mean S for the last five years:
m.S <- mean(S[(N+k-5):(N+k)])
```

```
  # Predict future remaining biomasses:
  for (i in (N+k+1):(N + (2*k))){
    mu.temp.rem[i] <- (m.S - ER) * (cR[i-1,1]+ Rem[i-1])
    Rem[i] ~ dnorm(mu.temp.rem[i], tau.process.rem)
  }
}
Data  :
list(
  ER = 0.346, # Exploitation rate
  N = 16,  # Number of years.
  k = 4,   # Number of back and forward predictions.
  mu.r = structure(.Data = c(
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
        3.6125,  4.0508,  4.5210,  4.7245,
        3.3631,  4.0320,  4.5093,  4.9283,
        3.2259,  4.7317,  5.0110,  5.2883,
        3.6764,  4.4880,  5.0677,  5.4712,
        3.7323,  4.9046,  5.4323,  5.7445,
        4.1910,  5.2932,  5.4855,  5.1118,
        4.0442,  5.1983,  5.3294,  4.9194,
        4.4239,  4.9570,  4.8074,  4.4546,
        4.0629,  4.7592,  4.3717,  4.1432,
        3.9924,  4.1718,  3.9054,  4.0013,
        3.6757,  4.0191,  3.8633,  4.0414,
        3.4455,  3.8165,  3.9915,  4.3835,
        3.0166,  3.7694,  4.2377,  4.4913,
        3.0035,  4.2698,  4.6895,  4.9431,
        3.3669,  4.4691,  4.5892,  4.5130,
        3.8843,  4.3837,  4.4578,  4.5608,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA),
   .Dim = c(24,4)),

  sigma.r = structure( .Data = c(
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
        0.1735,  0.1215,  0.1313,  0.1626,
        0.2532,  0.1620,  0.1253,  0.1423,
        0.2339,  0.1641,  0.1069,  0.1307,
        0.1313,  0.0938,  0.0899,  0.0896,
        0.1490,  0.1079,  0.0664,  0.0710,
        0.1085,  0.0845,  0.0746,  0.0967,
        0.1234,  0.0724,  0.0871,  0.1148,
        0.0826,  0.0702,  0.0757,  0.0908,
        0.0934,  0.0874,  0.0757,  0.0927,
        0.0822,  0.1221,  0.0683,  0.1041,
        0.0900,  0.1088,  0.1023,  0.1098,
        0.1119,  0.1304,  0.0836,  0.0788,
        0.0982,  0.1080,  0.0783,  0.0616,
        0.1386,  0.0868,  0.0799,  0.0552,
        0.1667,  0.0854,  0.0727,  0.0808,
        0.1170,  0.0988,  0.1127,  0.0899,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
            NA,      NA,      NA,      NA,
```

```
            NA,       NA,       NA,       NA),
    .Dim = c(24,4)),

mu.rem = c(NA,NA,NA,NA,27.6897,27.7751,31.177,9.97689,16.905,
13.0748,26.9193,19.7255,24.1398,19.9138,27.0248,21.0095,10.3992,
15.444,33.7682,26.0281,NA,NA,NA,NA),

sigma.rem = c(NA,NA,NA,NA,3.17082,3.83237,3.39582,1.98634,
2.42095,1.45758,3.181,2.15674,3.04264,2.0725,1.97866,1.65069,
1.01035,1.41315,2.98437,2.21879,NA,NA,NA,NA),

L = c(NA,NA,NA,NA, 17.655,13.864,15.517,19.184,18.513,26.178,
21.163,31.661,36.078,29.121,27.16,24.89,23.998,9.549,10.677,
21.963,NA,NA,NA,NA)
)
Inits :
list(
Rem = c(
  NA, NA,  NA,    NA,19.62,
25.86,31.36,13.68,20.99,14.68,
24.98,21.65,21.48,22.24,25.96,
20.87,12.66,15.34,32.92,25.93,
21.96,25.18,31.21,33.38),
S = c(
0.7804,0.6984,0.6193,0.7,0.6708,
0.687,0.6267,0.7718,0.6787,0.6653,
0.6118,0.5936,0.6644,0.6837,0.6549,
0.681,0.8065,0.6581,0.6644,0.6824,
0.6151,0.6439,0.668,0.7229),
alpha = c(
0.8183,0.9731,0.9653),
log.R = structure(.Data = c(
-197.3,-144.6,126.7,2.945,-118.3,
123.3,2.87,4.48,100.9,2.766,
4.293,4.299,2.292,4.285,4.217,
4.438,3.509,4.1,4.273,4.777,
3.329,3.945,4.622,5.11,3.238,
4.605,4.898,5.279,3.756,4.568,
5.103,5.552,3.74,4.981,5.342,
5.76,4.087,5.202,5.524,5.228,
4.261,5.329,5.147,4.942,4.357,
5.042,4.818,4.519,4.137,4.919,
4.358,4.122,4.014,4.24,3.918,
3.99,3.488,3.966,3.877,4.168,
3.244,3.75,4.006,4.39,3.05,
3.823,4.199,4.684,3.125,4.136,
4.567,4.889,3.377,4.527,4.591,
4.608,3.682,4.379,4.468,4.492,
3.605,4.418,4.336,4.633,3.601,
4.186,4.488,4.637,3.408,4.545,
4.501,4.332,3.733,4.472,4.205,
3.45),
.Dim = c(24,4)),
mu.S = 0.6696,
mu.r = structure(.Data = c(
  NA,  NA,  NA,   NA,-118.3,
123.4,2.88,4.64,100.9,2.975,
4.182,4.065,2.388,4.155,4.279,
4.357,   NA,    NA,    NA,    NA,
  NA,    NA,    NA,    NA,    NA,
  NA,    NA,    NA,    NA,    NA,
  NA,    NA,    NA,    NA,    NA,
```

```
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
   NA,   NA,   NA,   NA,   NA,
3.593,4.585,4.332,4.75,3.615,
4.439,4.59,4.771,3.336,4.478,
4.293,4.244,3.786,4.411,4.271,
3.493),
.Dim = c(24,4)),
mu.rem = c(
-16.29,   NA,   NA,   NA,    NA,
    NA,   NA,   NA,   NA,    NA,
    NA,   NA,   NA,   NA,    NA,
    NA,   NA,   NA,   NA,    NA,
    NA,   NA,   NA,   NA),
phi.alpha = 0.6412,
phi.beta = 1.728,
sigma.r = structure(.Data = c(
0.1489,0.1058,0.1179,0.109,0.1175,
0.2242,0.07687,0.1234,0.07398,0.1283,
0.1494,0.1288,0.1155,0.1869,0.05832,
0.1905,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
     NA,     NA,     NA,     NA,      NA,
0.05579,0.1493,0.08063,0.1326,0.1163,
0.2806,0.1204,0.1224,0.1618,0.108,
0.3158,0.1285,0.07708,0.08269,0.1101,
0.08468),
.Dim = c(24,4)),
sigma.rem = c(
3.79,5.821,3.197,1.73,   NA,
  NA,   NA,   NA,  NA,   NA,
  NA,   NA,   NA,  NA,   NA,
  NA,   NA,   NA,  NA,   NA,
2.569,5.565,3.733,7.215),
tau.S = 275.0,
tau.process.r = c(
3485.0,70.03,572.1,5.624),
tau.process.rem = 0.0116)
```

APPENDIX 2. ESTIMATION OF THE 2012 BIOMASS OF COMMERCIAL SNOW CRAB BY NICOLAS BEZ (JANUARY 2013).

DATA DECRIPTION

Densities (in weights) of commercial crabs were computed by dividing the weights by the swept areas. They are expressed in kg/m². Their spatial distribution covers the entire Gulf (Figure 1).

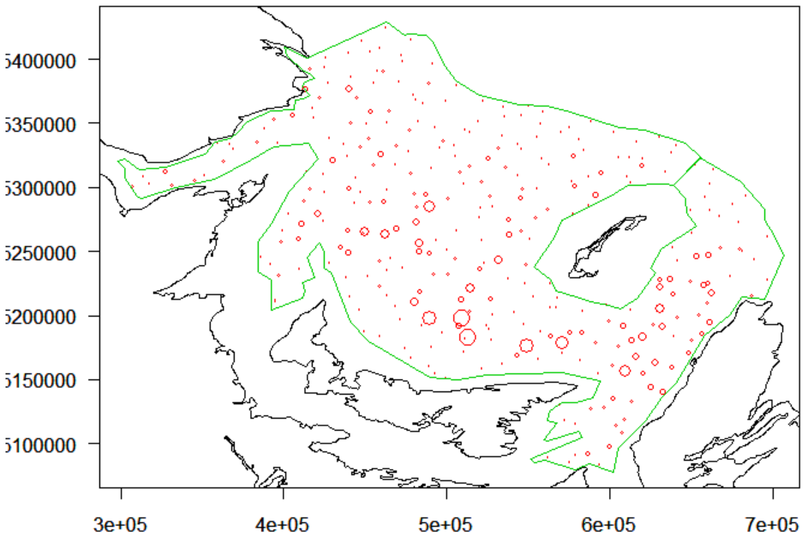


Figure 1: Spatial distribution of sample data in 2012 (N = 321 samples).

Sample coordinates and polygon vertices were projected in a consistent manner with the fine grid of bathymetry (type = UTM, zone = 20, data = NAD83). Measured depth at sample location was replaced by the interpolated bathymetry by migrating the nearest grid point to the sample points. So doing the depth measurements available at sample points and at interpolated grid nodes were consistent. Very small discrepancy existed between the two sets of depth measurement (see last year report). Over the 321 samples, 27% (N0=86) were null. The positive values ranged between 0.0001067 kg/m² and 0.01217 kg/m². The coefficient of variation was equal to 1.3.

Tow quality from 1 to 4 was distributed as follow : 214 samples were quality 1, 55 were quality 2, 41 were quality 3 and 11 were quality 4.

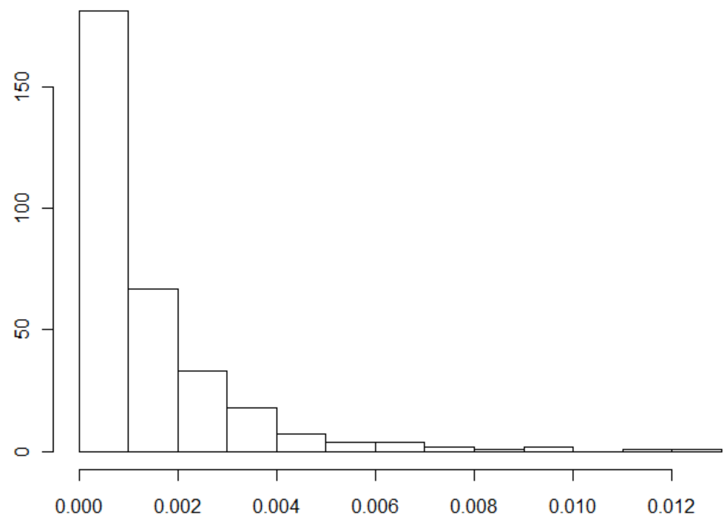


Figure 2: Histogram of mean weights (kg/m²).

STRUCTURAL ANALYSIS AND DATA QUALITY

Compared to all previous variograms (Figure 2), the 2012 variogram depicted similar characteristics except at the first distance lag. The first lag corresponds to samples that are very nearby in space. It was represented by 14 pairs of points 3,200 m apart on average, while the second lag was represented by 409 pairs of samples 10,000 m apart.

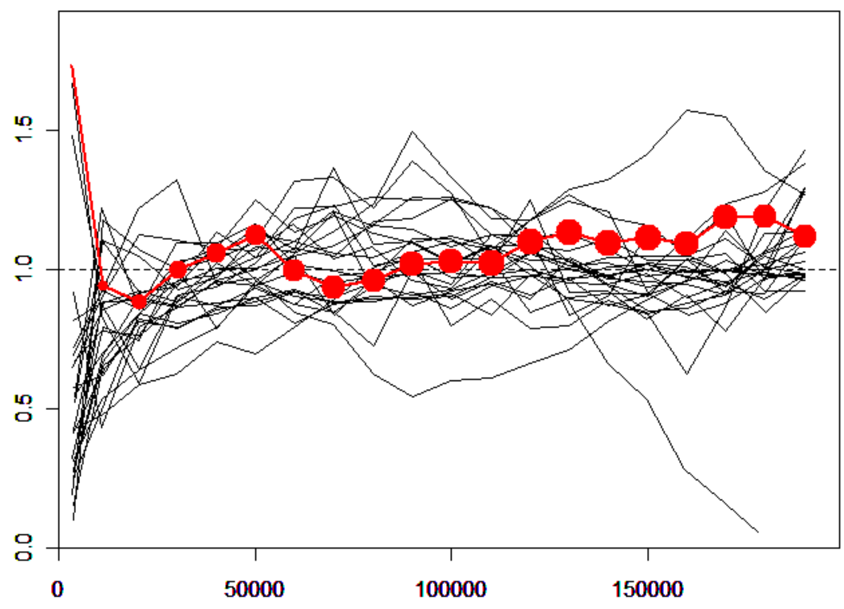


Figure 3. Standardised empirical annual variograms (omnidirectional) from 1988 to 2012. The 2012 variogram is in red with symbols proportional to the number of pairs of points available in each distance lag. Distances (x-axis) are in meters. Distance lag = 10,000m.

A detailed analysis of the first distance lag of the variogram indicated that three pairs of sample points were less than 2,000 m apart. These were the following samples :

	Pair 1		Pair 2		Pair 3	
Rank	123	124	141	146	268	269
Value (kg/m²)	0	0	0.00116	0.006	0.001155	0
Tow quality	1	1	2	4	4	1

Removing sample n°268 for instance, modified completely the first point of the variogram (Figure 4). Removing the 11 samples that were of bad quality (tow quality =4) also modified this first distance lag, the others being nearly not impacted. The first lag of the variogram was thus highly variable and uncertain. This is a common observation in applied geostatistics particularly with highly skewed distributions as this is the case here (Figure 2). This justifies the 0 weight given to the first distance lag in the model fitting procedure.

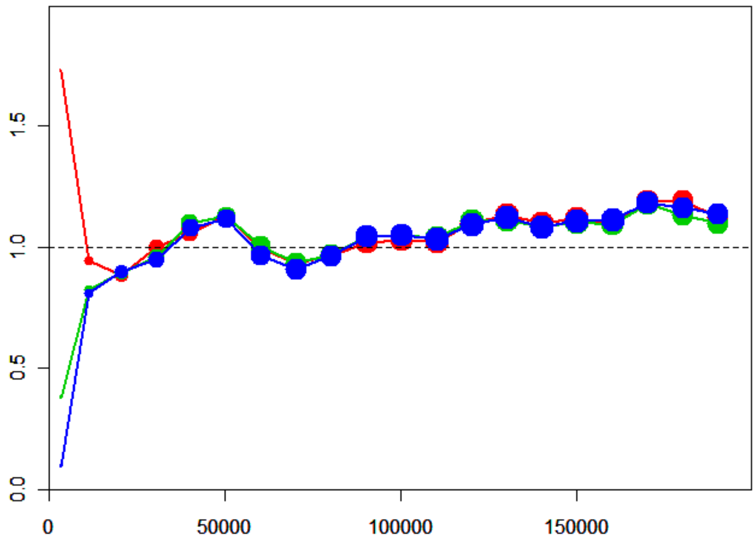


Figure 4: Standardised empirical annual variograms (omnidirectional) for 2012. Impact of data selection on the first distance lag : in green result without sample n° 268 ; in blue result without tows of bad quality (tows with tow quality = 4). Symbols proportional to the number of pairs of points available in each distance lag. Distances (x-axis) are in meters. Distance lag = 10,000m.

Depth

The general pattern of the relationship between crab density and depth is observed in 2012 : it increases rapidly after 50 m, flattens and then reduces to 0 after 200 m (Figure 5).

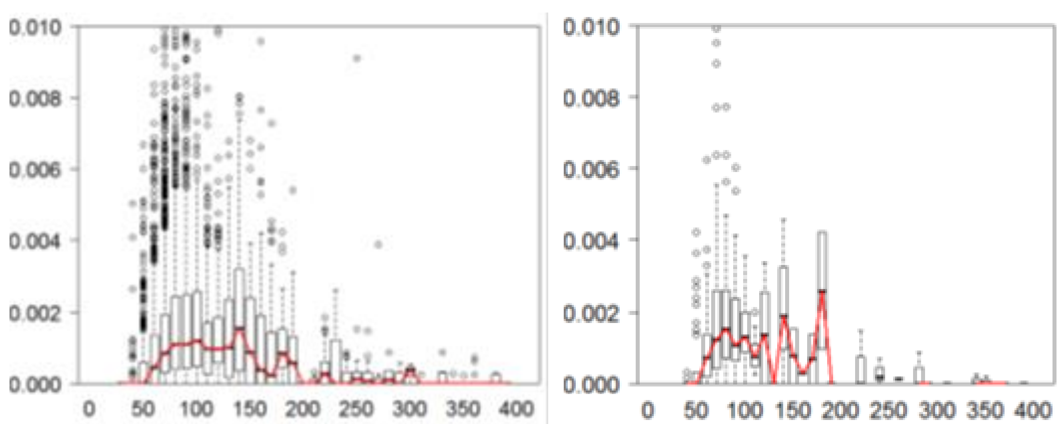


Figure 5: Relation with depth. Left : all year. Right : 2012 only.

CROSS VALIDATION AND CHOICE OF A VARIOGRAM MODEL

Two different variograms were computed. The annual variogram using 2012 data only with tows of bad quality exlcuded, and the mean variogram over 2010-2012 (Figure 6). This latter variogram was obtained by first computing annual variogram standardised by the annual variances. The standardisation of the annual variogram avoids that one particularly heterogeneous variable drive the final result. The three variograms were then averaged (with weights equal to the number of pairs of points behind each variogram value). Finally the averaged variogram was re-scaled to the 2012 experimental variance.

Models were chosen to be a combination of a nugget effect, a sherical model and a linear model. The estimation of the parameters of the model was done by automatic fitting based on a least squares algorithm considering the numbers of pairs.

The weight of first lag of the variogram was set to 0 to avoid considering this point which was considered not enough robust.

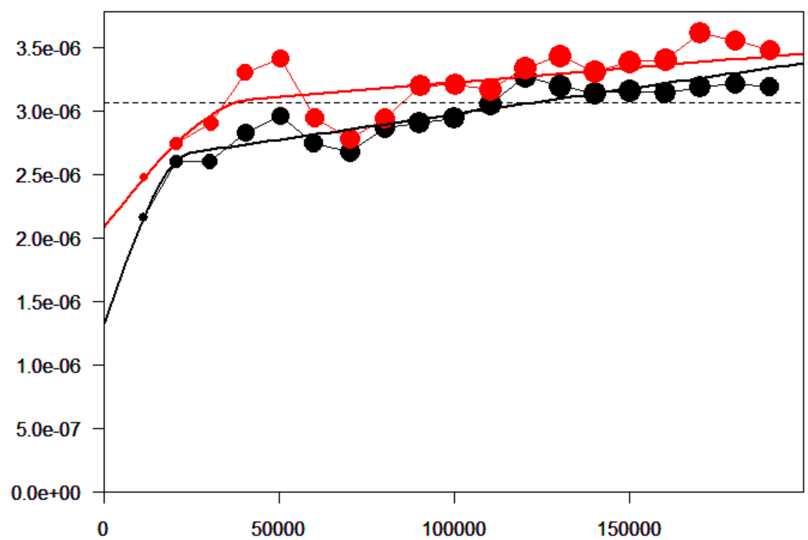


Figure 6: Annual variogram without tows whose quality was registered equal to 4 (red) and 2010-2012 mean variogram (black). The first lags of each variogram have been down weighted to 0. Distance lag = 10,000 m. Distances are in metres. Circles are proportional to the number of pairs of poitns participating to the computation. Variogram models are overlayed.

	Nugget effect	Spheric		Linear
	Sill (kg^2/m^4)	Sill (kg^2/m^4)	Range (m)	slope
Annual variogram	2.09e-6	9.09e-7	42,192	2.279e-12
Mean variogram	1.32e-6	1.25e-6	24,698	4.036e-12

Despite the strong similarity between variograms, cross validations were done by removing one by one each sample and by re-estimating it by the standard procedure (ie Kriging with External Drift with a neighbourhood of 32 samples over 8 angular sectors). This was also done with ordinary kriging in order to compare results not only between models but alos between estimation procedures.

Statistical characteristics of the residuals are the same (Figures 7-9). No bias in both cases, same shift of the median with regards to the mean and same dispersion around the mean. Choosing on the basis of the performance is thus impossible; the two models behave similarly.

As expected, the largest residuals were associated with the largest densities. No general trend in the performance between the two models appeared in this matter either. However for large densities (i.e. densities larger than 0.005 kg/m²), the leave one out procedure performed better with the annual variogram rather than the triannual mean variogram (Figure 8). However the difference was one order smaller than the targeted density, i.e. the estimate obtained with the annual variogram was 0.00025 larger than the one obtained with the triannual variogram.

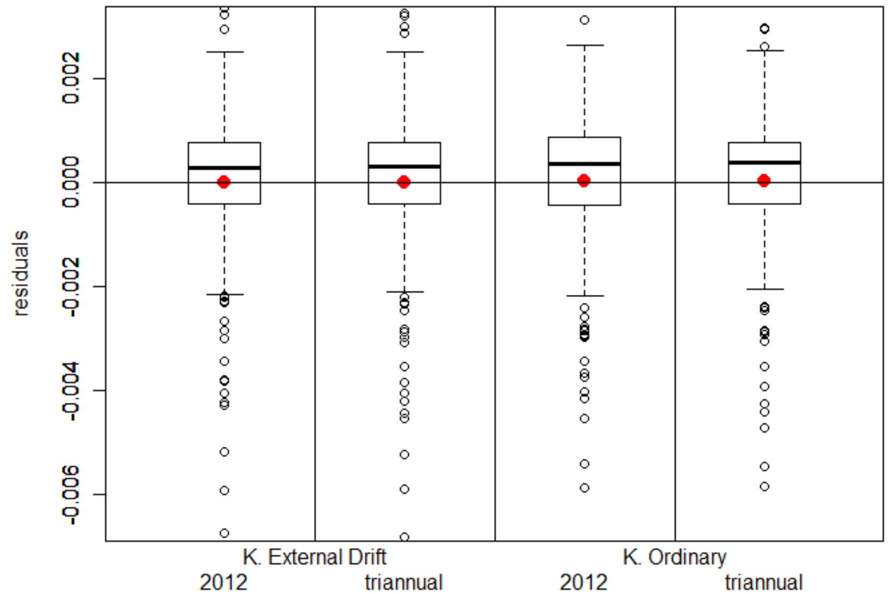


Figure 7: Cross validation. Boxplot of residuals (Z-Z) for each model and for KED or KO. The horizontal line represents the 0 line. Means residual are represented by red points.*

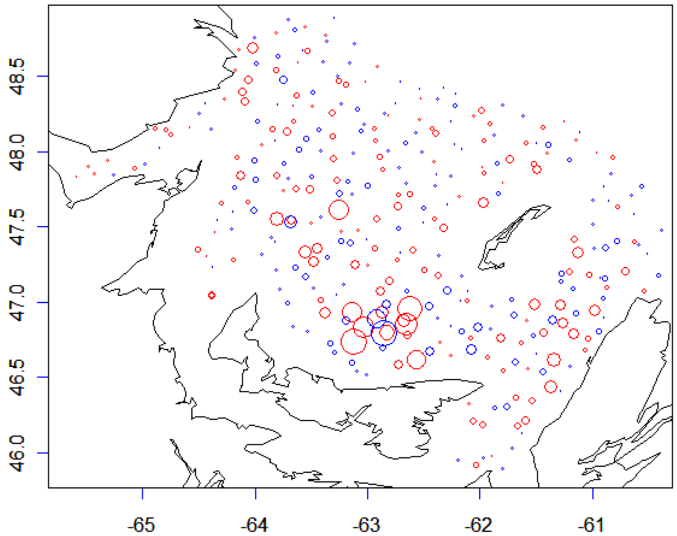


Figure 8:. Cross validation. Spatial distribution of the differences between cross validation estimates with KED : estimates from annual variogram minus estimates from triannual mean variogram. In red, positive differences (annual variogram > mean variogram). In blue, negative differences (annual variogram < mean variogram).

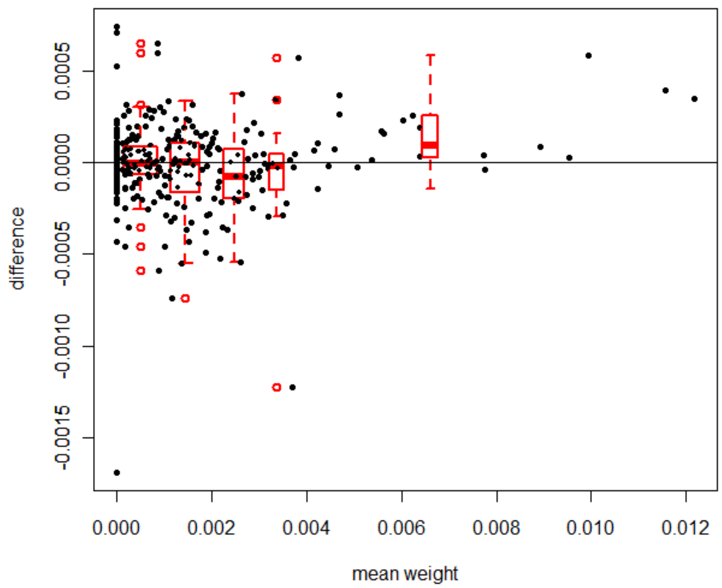
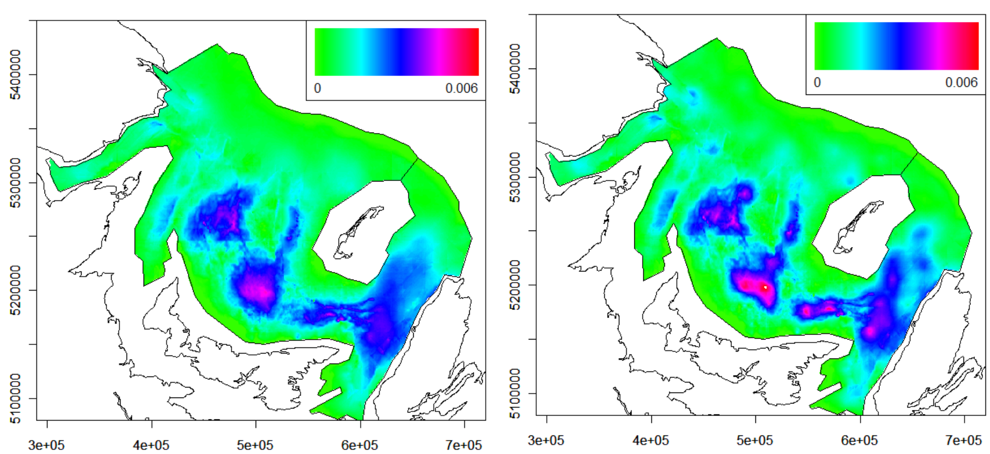


Figure 9: Cross validation. Differences between re-estimates (estimates with the “annual” minus estimates with the “mean” variogram) as a function of the observed value. Positive y-axis represents estimates which are larger when performed with the annual variogram. Reverse for negative part.

PUNCTUAL ESTIMATIONS

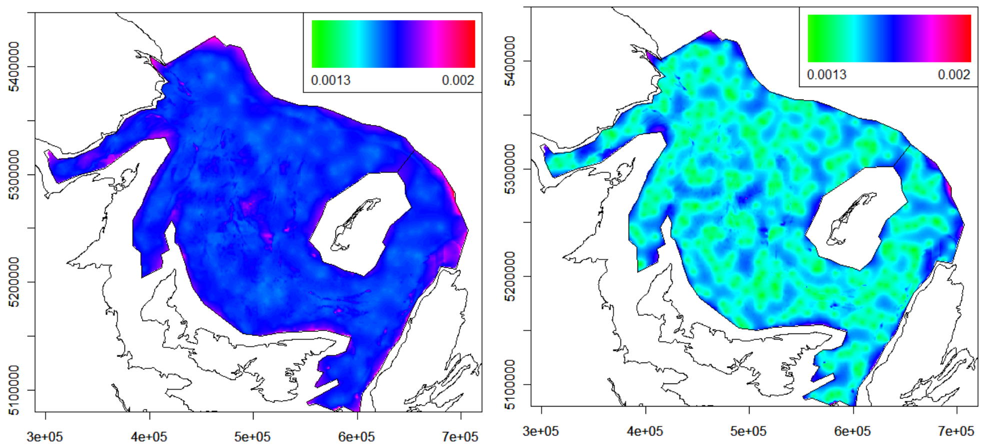
Interpolations were done in the projected space using the regular bathymetric grid for the interpolation. Kriging with external drift (Figure 10) and ordinary kriging (Figure 11) were performed with the two models over the targeted 20-200 polygon. The moving neighborhood was made of 32 points within a circle of 125 km radius with 8 angular sectors. Cross validation indicated that all the negative estimations were associated to null data. We thus attributed a 0 density value to the negative interpolated values.



Annual 2012 variogram

Mean Triannual variogram

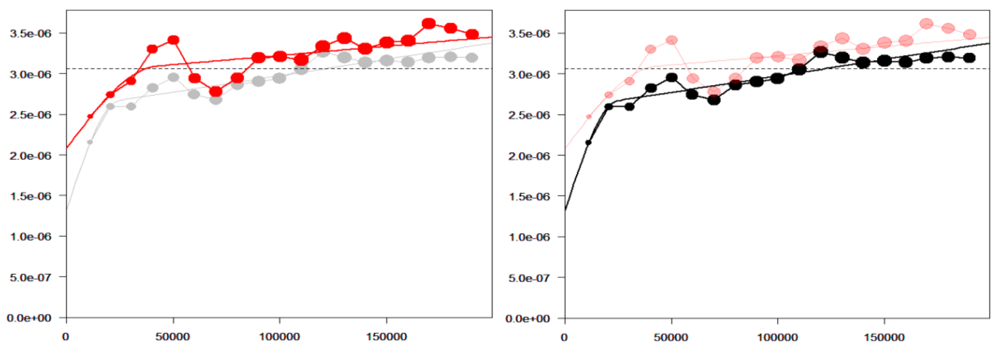
Figure10a: Kriging with external drift (kg/m²).



Annual 2012 variogram

Mean Triannual variogram

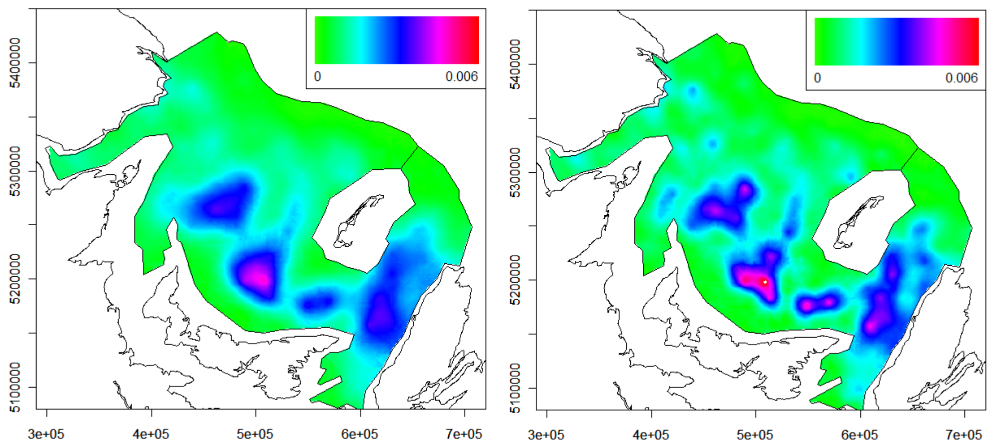
Figure 10b: Kriging with external drift. Standard deviation of the estimation.



Annual 2012 variogram

Mean Triannual variogram

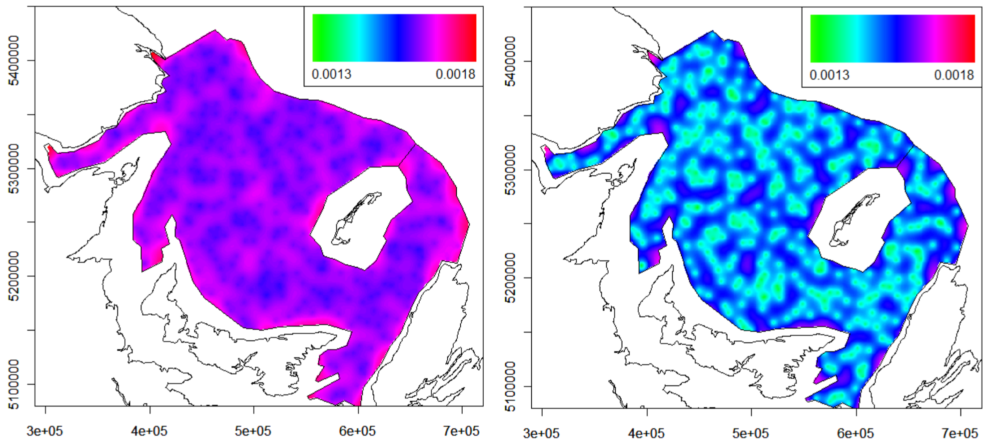
Figure 10c: Variogram model used for the interpolation.



Annual 2012 variogram

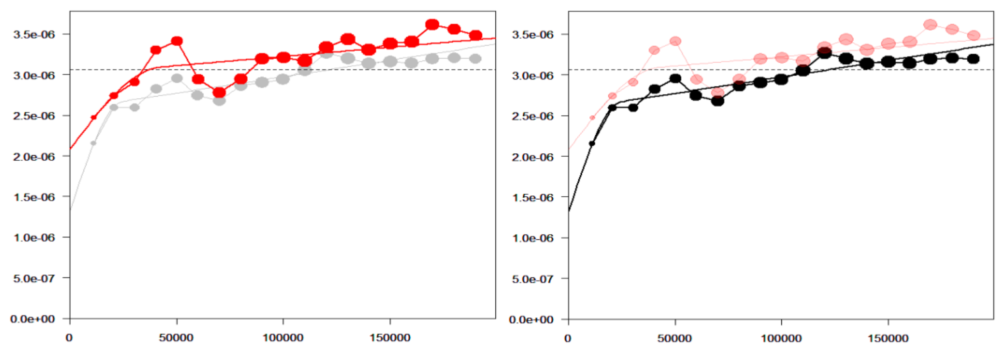
Mean Triannual variogram

Figure 11a: Ordinary Kriging.



Annual 2012 variogram Mean Triannual variogram

Figure 11b: Ordinary Kriging. Standard deviation of the estimation.



Annual 2012 variogram Mean Triannual variogram

Figure 11c: Variogram model used for the interpolation.

As expected :

- krigings with the mean triannual variograms led to maps with more details as the variogram gets a smaller proportion of nugget, and
- krigings with the triannual variograms led to smaller standard deviation as the variogram was i) smaller and ii) gets less proportion of nugget.

GLOBAL ESTIMATIONS

After projection (UTM, 20, NAD83), polygons were 58 047km². The estimations of snow crab biomass by averaging punctual estimations over the 20-200 polygon were the following :

	Surface (km²)	Arithmetic total (tonnes)	Mean of the punctual KED (tonnes)	Mean of the ponctual KO (tonnes)
2012	58,047	78,038	75,082	77,785
2010-2012			73,849	76,591

The differences coming from the choice of the model (annual versus triannual variogram) were small compared to the differences coming from the choice of the method (KED vs KO). Using depth as an external factor (ie a factor that drives the variations of crab density) probably tends to reduce the extrapolation of crab density in the border of the polygon were depth is (very) large.

Polygon kriging can not be performed with KED but with KO. Coefficient of variation were obtained by a discretization of the polygon with 200 grid mesh in both dimensions (longitude and latitude).

	Surface (km^2)	Arithmetic total (tonnes)	Global estimation KO (tonnes)	CV (%)
2012	58,047	78,038	77,926	7.2%
2010-2012			77,596	6.6%

GENERAL COMMENTS

Some differences exist between the annual variogram and the triannual average variogram without impacting significantly the final results.

The use of the samples with low tow quality should be addressed at some point (for the variogram calculation, for the biomass estimation).