A geostatistical method for assessing biomass of tuna aggregations around moored Fish Aggregating Devices with star acoustic surveys

Doray M.¹, Petitgas P.², Josse E.³

¹IRD/Ifremer, US acoustique halieutique, centre IRD de Bretagne, BP 70, 29280, Plouzané, France
²Ifremer, EMH, rue de l’Île d’Yeu B.P. 21105 44311 Nantes Cedex 03 France.
³IRD, US acoustique halieutique, centre IRD de Bretagne, BP 70, 29280 Plouzané, France.

*: Corresponding author: Mathieu Doray: Ifremer, EMH, rue de l’Île d’Yeu B.P. 21105 44311 Nantes Cedex 03 France, phone: 33.2.40.37.40.00 , fax: 33.2.40.37.40.01 , email address: mathieu.doray@gmail.com

Abstract:

Universal kriging was used to model the spatio-temporal variability in the acoustic density of tuna aggregations recorded during star echosounding surveys around moored Fish Aggregating Devices (FAD) in Martinique (Lesser Antilles). The large-scale deterministic drift in the tuna spatial distribution was modeled using an advection-diffusion equation applied to animal grouping. Residuals from the drift were modeled as a random component with small-scale spatial correlation. An estimation variance formula was derived from this deterministic/statistical mixed model to assess the mean precision of density estimates of daytime tuna aggregation. The mean relative error obtained with our star design for daytime surveys was 24%. The methodology was applied to estimate daily maxima of tuna biomass around moored FADs during 4 monthly sea cruises. The daily peak of tuna biomass aggregated around moored FADs was 9 MT on average (SD = 4). Estimation variances for different survey designs were compared for optimizing sampling strategy.

Keywords: geostatistics, tuna, moored fish aggregating devices, advection-diffusion model, acoustic survey precision
Introduction

Fisheries operating around moored or drifting Fish Aggregating Devices (FADs) have been developing since the 1980's (Fonteneau et al. 2000; Fréon and Dagorn 2000). Currently, they provide about half of worldwide tuna catches (Fonteneau et al. 2000), taking advantage of the peculiar aggregative behavior of large pelagic fish around floating objects. In spite of the economical importance of fish aggregations around FADs, quantitative knowledge about these shoals is scarce. Direct survey methods, such as echosounding star surveys, were developed to map fish density (Josse et al. 2000) and to characterize fish aggregations around moored FADs (Doray et al. 2006, 2007). Nonetheless, the abundance of fish aggregated around a FAD at a given time has never been precisely estimated. We have designed a geostatistical methodology to estimate the mean density of tuna around moored FADs along with the precision of that estimation (see Chiles and Delfiner (1999) for a comprehensive textbook on geostatistics).

Computing precise density estimates based on acoustic data collected during star surveys around a FAD involves i) taking into account the movements of the floating object ii) correcting the oversampling of the center of the sampled area (Doonan et al. 2003), and iii) coping with the high temporal variability of tuna abundance around the moored FAD (Doray et al. 2006). To account for movement of the FAD, we referenced geographical positions of samples to known (recorded) FAD positions. We used a universal kriging model similar to that of Petitgas (1997) to model the time-invariant spatial structure of aggregations and to correct for oversampling the center of the survey area in the kriging process.

In this paper we analyze the spatial distribution of daytime sub-surface tuna aggregations as observed
The mean daytime spatial distribution of tuna was estimated by averaging their acoustic density in time using all surveys. This consistent component in the tuna spatial distribution (later referred to as drift or trend) was modeled in a deterministic advection-diffusion framework. Then the residuals from the drift were modeled as a random component with spatial correlation. The average spatial distribution of tuna aggregations was thus modeled as the sum of a deterministic component for the drift and a statistical one for the residuals. An estimation variance equation was derived from this mixed model for estimating the mean precision of daytime acoustic star surveys. Abundance estimates and estimation variances were also computed for individual surveys. Estimation variances of density obtained with our star survey and other survey designs were compared in order to optimize sampling strategy.

Material and methods

Data

From January, 2003, to April, 2004, repeated acoustic observations were undertaken around the island of Martinique on monthly sea cruises on board the 12 m chartered vessel “Béryx” which was equipped with a Simrad EK60 scientific echosounder (version 1.4.6.72) connected to two hull-mounted, spherical split-beam transducers (ES38-B and ES120-7G), operating vertically at 38 and 120 kHz frequencies. Each transducer had $7^\circ$ beam angles at -3 dB, with pulse lengths set to 0.512 ms for both. In situ on-axis calibration of the echosounder was performed before each cruise using standard methodology (Foote 1982). Acoustic data were archived in the international hydro-acoustic data format (HAC) at a $-80$ dB threshold with Movies+ software (Weill et al., 1993). A total of 516 star acoustic surveys conducted around moored FADs in Martinique (Lesser Antilles).
surveys (Josse et al. 2000) (Fig. 1a) were conducted within a radius of 400 m around 2 FADs moored at 2000 and 2500 m depths, on the leeward coast of Martinique at 7 (‘coastal FAD’) and 25 (‘offshore FAD’) nautical miles from the coast, respectively (Doray et al. 2006). During daylight hours, star surveys were conducted around each FAD approximately every 2 hours, yielding an average of 11 star surveys per cruise around each FAD. A unique ‘sub-surface’ tuna aggregation was observed in the vicinity of both FADs from 35 to 87 m depths during almost all daytime surveys. The daytime spatial distribution of the aggregation was assumed to display enough stability to allow for the estimation of the tuna biomass around the FADs in each single acoustic star survey. In-situ Target Strength (TS) of sub-surface tuna was analysed from the acoustic data to estimate the mean acoustic response of a single tuna (Doray et al. 2006). In addition, fishing experiments were conducted to assess the mean weight of sub-surface tunas around the FADs (Doray et al. 2007).

The echograms were processed with an image analysis algorithm implemented in Movies+ (Weill et al. 1993) to extract the tuna acoustic shoals from the dense surrounding scattering layers (see Doray et al. 2006 for details). This algorithm was applied to data collected at the 120 kHz operating frequency in 50 daytime surveys, conducted from April to August, 2003. The acoustic records along the survey track were binned in Elementary Sampling Units (ESU) of about 15 m in length. Acoustic densities (area backscattering coefficient: $s_a$ (Simmonds and MacLennan 2005)) of tuna shoals detected from 0 to 200 m depth were averaged within each ESU and allocated to the geographical position of the ESU center.

A schematic outline of the methodology designed to estimate the mean density of tuna aggregation and the precision of star acoustic surveys is presented in Fig. 2. The position of an ESU sampled at time $t$ was first referenced to the position of the head of the FAD at that time, which moved over hundreds of
106 meters during a star survey depending on the current. Referencing the ESU positions to the FAD position hence allowed restoring the geometry of star acoustic surveys conducted in strong currents. During the surveys, FAD positions were recorded each time the vessel passed near the device. The time varying FAD positions were estimated by modelling the trajectory of the FAD during a survey as a function of time, based on recorded FAD positions. An example map with samples referenced to the FAD position is illustrated in Fig. 3.

Once referenced to the FAD head, the ESU positions were further standardised, following the procedure of Okubo and Chiang (1974). The ESU coordinates in each survey were referenced to the gravity center of non-zero values and their standard deviation was adjusted to equal the average standard deviation of the tuna aggregation coordinates computed over all surveys. This procedure resulted in scaling the dimensions of the fish aggregation around the FADs in each individual survey to that of the mean aggregation estimates over all surveys. Such scaling was deemed realistic because the tuna aggregation appeared to be relatively isotropic and stable from one survey to another. However, fluctuations of the total tuna abundance, aggregative behavior at the tens of meter scale, and environmental parameters (e.g. currents) induced some variation in the anisotropy and spreading of the aggregation. Thus, we needed to filter out these variations (residuals) to characterize the mean spatial distribution (drift) of the aggregation.

GPS positioning error (about 7 m in our case) was significant at the scale of our study. For each survey, the surveyed area was gridded into 15 m x 15 m cells and the average of the samples in each cell was calculated and attributed to the position at the center of the cell. This procedure allowed for smoothing out of GPS positioning errors by insuring that the center of each cell was located inside the 7 m circle.
of uncertainty surrounding its GPS position. Averaging tuna acoustic densities within each cell also
reduced the spatial variability induced by fish movements at the tens of meter scale during the surveys.

**The space-time model**

The spatial distribution of tuna around a FAD can be viewed as an outcome of their aggregative
behavior in the vicinity of the device, i.e. an emergent property (Allen and Starr 1982; Deneubourg and
Goss 1989; Parrish et al. 2002) resulting from density dependent processes operating at the scale of
individual tunas (Parrish and Edelstein-Keshet 1999). Within this framework, spatial distributions of
tuna observed during repeated surveys should be relatively similar, provided that i) the same density
dependent processes apply, and ii) environmental conditions are comparable. Our observations indicate
that the overall shape of the density surface of the tuna aggregation around each FAD stayed relatively
stable from one survey to another. We therefore modeled the density surface using a universal kriging
approach (Matheron 1971; Petitgas 2001), which allows one to distinguish between a time invariant
drift and spatially correlated residuals. Our model was adapted from a space-time model used to
estimate sole (*Solea solea*, L.) egg mean density in a nursery ground (Petitgas 1997).

The model was defined within a circular area $V$ of radius 160 m in standardised coordinates, centered
around the gravity center of the aggregation and over which 95% of the mean acoustic energy of the
tuna aggregation was recorded.

Let us denote the tuna acoustic density in cell $k$ and time $t$ as the realization $Z(k,t)$ of a random function
$Z(x)$. $Z_V(t)$ denotes the average tuna density at time $t$ on the area $V$. The relative density at $k$ and $t$ is the
contribution of cell $k$ to $Z_V(t)$ and is denoted $Z_p(k,t)$. The space time model then reads
We have derived the model estimating $Z_p(k,t)$,

$$Z_p^*(k,t) = \frac{Z(k,t)}{Z_a^*(t)},$$

where $Z_a^*(t) = \frac{1}{N_s(t)} \sum_i Z(k_i,t)$ is the arithmetic mean of the $N_s(t)$ acoustic densities sampled during survey $t$. Our intention was not to precisely estimate $Z_v(t)$ with $Z_a^*(t)$. $Z_a^*(t)$ was solely used to scale acoustic densities observed during surveys with different global abundances.

The mean spatial correlation structure in the residuals was modelled by a variogram which was subsequently used to compute the estimation variance of the mean abundance estimate of i) any daytime survey, ii) abundance maxima, and iii) other survey designs (Fig. 2). The model was structured so that estimates from each of the surveys were interpreted as repeated estimates of the same phenomenon over time. In model terms, this meant that averaging over time amounted to taking the expected value of the random function $Z_p(k,t)$. The estimation for a given survey grid was then performed using data from that grid only and inference of the spatio-temporal structure in the residuals was not necessary. Hence, we solely modeled the average in time of the spatial structure in the residuals. The model considered is thus a spatial model whose inference implies the use of several
realizations of an underlying phenomenon.

**Estimation of the drift using advection-diffusion modeling of aggregations**

Matheron (1971) demonstrated that given one survey, drift and process residuals cannot be estimated together using the same data. We therefore took advantage of the repeated surveys to estimate $m(k)$.

We first computed the two-dimensional (2D) arithmetic mean density surface of tuna $m_s^*(k)$:

\[
(3) \quad m_s^*(k) = \frac{1}{N_s} \sum_{t} Z_p^*(k,t),
\]

where $N_s$ is the number of times cells $k$ was sampled across all surveys. This surface was centered around the gravity center of the aggregation and was remarkably isotropic, with a dome shape and truncated tails near the aggregation boundaries (Fig. 4). In this regard, the mean tuna spatial distribution interestingly resembled those of midge swarms modeled by Okubo and Chiang (1974) and Okubo et al. (2001) using advection-diffusion equations. We therefore used these equations to model $m_s^*(k)$.

Okubo et al. (2001)’s model formulates that an aggregation of animals results from animal movements due to different causes and occurring in opposite directions. The model assumes that, within an animal group, the net one-dimensional (1D) flux $J$ of individuals through a vertical section of the group results from a flux due to diffusion and an opposite flux due to advection (Okubo et al. 2001):

\[
(4) \quad J = uS - D \frac{\partial S}{\partial x},
\]
where $S$ is the animal density at point $x$, while $D$ and $u$ are the diffusion and advection coefficients, respectively.

When diffusion equals advection, $J = 0$, a stable group is maintained, and (4) yields:

\[
\frac{S}{\partial x} = \frac{D}{u}
\]

Following Okubo and Chiang (1974), a smoothing procedure (kernel smoothing) was used to formulate the drift in an isotropic and continuous manner, as a function of the distance $r$ to the gravity center of the aggregation. That procedure facilitated the subsequent fit of the kernel-estimated surface with the advection-diffusion model.

Let $l$ denote the tolerance used to define distance classes. In this case, the width of distance classes is defined as $r + l - (r - l) = 2l$. The kernel estimate of the relative tuna density $m^*_k(r)$ in the distance class $r$ was computed as:

\[
m^*_k(r) = \sum_{r-l \leq d(x) < r+l} \frac{k_h(d(k) - r)}{k_h(d(k) - r)} \times m^*_a(k)
\]

where $d(k)$ is the distance from the center of cell $k$ to the aggregation gravity center and $k_h$ is the Epanečnikov kernel of bandwidth $h$ (Stoyan and Stoyan 1994).

Applying Okubo et al. (2001)’s model with our tuna model notations, (5) writes:
Let $ADR$ denote the advection-diffusion ratio $D/|\mu|$. Similarly to Okubo and Chiang (1974), $ADR$ varied with fish density $m_k^*(r)$. That relationship was modeled explicitly, resulting in an advection-diffusion model-based estimate $m_{mod}$ of the drift, expressed as a function of $r$.

### Spatial correlation in the residuals

Spatial correlation modeling

Residuals for a survey $t$ were calculated relative to the advection-diffusion model:

$$R^*(k,t) = Z^*_k(k,t) - m_{mod}(k)$$

The advection-diffusion model was applied for each grid cell $k$.

We modeled the mean correlation structure across all surveys with a variogram $\gamma$. The experimental variogram in each survey $\gamma^*(h)$ was estimated on the residuals $R^*(k,t)$. Then, these variograms were averaged and the result was modeled by the sum of a nugget effect and a spherical model. The experimental variogram of the residuals for survey $t$ was estimated using a weighting of the samples (Rivoirard et al. 2000):
\[
\gamma_i^*(h) = 0.5 \sum_{x_i-x_j=h} a(k_i,t)a(k_j,t) \left[R^*(k_i,t) - R^*(k_j,t)\right]^2
\]

where \( a(k,t) \) is the area of influence of each sampled cell \( k \). Areas of influence were calculated by finely discretizing the study area \( V \) and applying the following formula within circular areas of radius 50 m centered around sampled cells centers:

\[
a(k,t) = \frac{n_a}{N_c}
\]

where \( N_c \) is the total number of small cells in the circular area and \( n_a \) is the number of small cells nearest to the sampled cell center. Weighting by area of influence was introduced to deal with the sampling heterogeneity of the star survey pattern when estimating the variogram.

The average in time of the residual variogram was estimated using

\[
\gamma^*(h) = \frac{1}{N_s} \sum_t \gamma_t^*(h),
\]

where \( N_s \) is the number of surveys considered.

Control of the average model fit and for particular surveys

The modeled residual variance should be comparable to those of the empirical residuals, \( R^*(k,t) \). This was checked on average over all surveys (i.e., realizations) as well as for particular surveys. The average variance in the residuals was estimated empirically as well as with the model and the values
compared. The empirical estimate, \( s^2_{R}(V) \), was the average over all surveys of the survey specific experimental residual variances \( s^2_{R}(t) \). The model based estimate was

\[
\tilde{\gamma}_{VV} = \frac{1}{V^2} \int_{V} \int_{V} \gamma(|x - y|)dy ,
\]

where \( \gamma \) is the mean variogram model of residuals and \( x \) and \( y \) are coordinates of cells within the estimation domain \( V \).

For a particular survey \( t \), if the drift model \( m_{\text{mod}}(k) \) was not adequate, the empirical residual variance for that survey \( s^2_{R}(t) \) would be inflated. The ability of the residual average variogram model \( \gamma \) to estimate the empirical residual variance of single surveys was used as a quality control parameter of the adequacy of the spatial model fit to particular surveys. The adequacy for each survey \( t \) of the drift model \( m_{\text{mod}}(k) \) and of the residual model \( \gamma \) was assessed by the difference \( D_{R}(t) \) between the empirical residual variance \( s^2_{R}(t) \) and the model-based variance \( \tilde{\gamma}_{VV} \). A low \( D_{R} \) for a survey \( t \) indicated a good fit of the model to the data for that survey. Studying the distribution of \( D_{R}(t) \) allowed for the assessment of the adequacy of the spatial model to all surveys.

**Aggregation density and estimation variance**

**Mean precision of star surveys**

For any survey \( t \), the areal mean of the tuna density \( Z^*(t) \) was estimated by kriging (Matheron 1971).
The kriging estimate of the areal mean is

\[ Z^*_y(t) = \sum_i \lambda_i Z(k_i, t), \]

where \( \lambda_i \) are kriging weights summing to unity. Kriging weights are defined so as to minimize the estimation variance (Rivoirard et al. 2000). The estimation variance is not conditional on the data values of a given survey and writes (Petitgas, 1997):

\[ \sigma^2_E = E\left[ Z^2_y(t) \right] \times Var\left( Z_{p,y}^* (t) - Z_{p,y}^* (t) \right) = E\left[ Z^2_y (t) \right] \times \sigma^2_{spa}, \]

where \( Z_{p,y}^* (t) \) is the estimation of the areal mean of the relative density in survey \( t \) and \( \sigma^2_{spa} \) the estimation variance of \( Z_{p,y}^* (t) \).

Within the framework of the universal kriging model, \( \sigma^2_{spa} \) develops in two error variance terms, \( \sigma^2_m \) and \( \sigma^2_R \), which are associated with the drift and the residuals respectively (Matheron 1971)

\[ \sigma^2_{spa} = \sigma^2_m + \sigma^2_R = \frac{\sigma^2_E}{E\left[ Z^2_y (t) \right]}, \]

\( \sigma^2_m \) was estimated by 1- \( R^2 \), \( R^2 \) being the coefficient of determination of the fit of \( m_k^* (k) \) with \( m_{mod} (k) \) (Scherrer 1984).
\[ R^2 = \frac{\sum (m_{\text{mod}} - \overline{m}_k)^2}{\sum (m_k - \overline{m}_k)^2} \]  

The error term associated with the residuals was estimated by a geostatistical estimation variance (Rivoirard et al. 2000)

\[ \sigma_R^2 = 2\sum_i \lambda_i \gamma(x_i, V) - \overline{\gamma}(V, V) - \sum_i \sum_j \lambda_i \lambda_j \gamma(k_i, k_j) , \]

where \( \gamma \) is the model of the mean variogram of residuals, \( \lambda \) kriging weights summing to unity and \( k \) the cells sampled by the most regular star survey conducted during the cruises.

\[ E[Z_v^2(t)] \] was estimated by \( E[Z_a^* (t)] \), the time average of squared daytime arithmetical mean acoustic densities \( Z_a^*(t) \) and the standard (estimation) error \( SE \) by

\[ SE = \frac{\sigma_E}{E[Z_a^* (t)]} \]

Abundance estimates

The spatial mean of the tuna aggregation acoustic density, \( Z_v(t) \), was estimated for each survey \( t \) by kriging the survey mean relative density surface, \( Z_p(k, t) \), then scaling the kriged relative estimate with the survey data average, \( Z_a^*(t) \). For kriging, we used the time-invariant model \( \gamma \) of the mean residual
variogram; the survey mean estimate was calculated as

\[
Z^*_v(t) = \sum_i \lambda_i Z(k_i, t) = \sum_i \lambda_i Z_p(k_i, t),
\]

where \( \lambda_i \) are the kriging weights summing to unity.

Because of the higher density sampling near the FAD centers introduced by the star survey geometry, we used kriging to optimally weight the sample values according to each survey configuration and average variogram structure.

Finally, the abundance \( A \) and biomass \( B \) of tuna in the area \( V \) were computed as:

\[
A(t) = \rho_A \times V = \frac{Z^*_v(t)}{\sigma_{bs}} \times V \quad \text{and} \quad B(t) = A(t) \times W,
\]

where \( \rho_A \) is the density of tuna within the aggregation (Nb. fish.m\(^{-2}\)), while \( \sigma_{bs} \) and \( W \) are the mean backscattering cross-section (Simmonds and MacLennan 2005) and mean weight of a single tuna, respectively. The mean TS value of a sub-surface tuna given by Doray et al (2007) (-35 dB) was used to compute \( \sigma_{bs} \) as: \( \sigma_{bs} = 10^{\frac{TS}{10}} \). We used 2.7 kg as \( W \) estimate, based on the mean fork length of sub-surface tuna given by Doray et al (2007) (58 cm) and a length-weight relationship established for blackfin tuna caught around moored FADs in Martinique (Rivolaen et al 2007).

This methodology was applied to assess daily maxima of tuna abundance observed during the sea cruises. Daily maxima of abundance were defined by selecting the survey with the highest mean
282 arithmetic tuna density ($Z_a(t)$) in each surveyed 24-h period.

283 Other survey patterns

284 Once spatial correlation in fish distribution is modeled, one can compute and compare estimation
285 variances of various survey designs (Petitgas and Lafont 1997). We compared the precision in our star
286 survey with that for other star survey designs. We assumed that all surveys would provide similar $m$,
287 $\sigma_m^2$ and $\gamma$ and focused on the differences in $\sigma_R^2$ due to differences in survey design.

288 The sample positions needed to compute estimation variances were obtained by applying various
289 survey designs on a virtual tuna aggregation. We positioned a circular aggregation in a 400 m x 400 m
290 area divided into 15 x 15 m cells centered around a virtual FAD. Aggregation dimensions and the
291 relative position from the FAD were identical to those of observed average sub-surface tuna
292 aggregation (Doray et al. 2006). Survey designs described in Table 1 were applied to the virtual
293 aggregation. Branches of star designs were sampled along two parallel tracks ('with duplicate' as in Fig.
294 1a) or along a single track ('without duplicate' as in Fig. 1b). Sample positions obtained with each
295 design were used to compute residual estimation variances $\sigma_R^2$ in a circular estimation area of radius
296 160 m, centered around the gravity center of the aggregation. Estimation variances were computed
297 using equation 17 and the variogram model $\gamma$.

298 Statistics were implemented using the R language (R Development Core Team 2007). Geostatistical
299 computations were implemented using EVA2 software (Petitgas and Lafont 1997) and R packages
300 geoR (Ribeiro and Diggle 2001) and RgeoS (Renard and Bez 2005).
301 Results

302 Drift characterisation

As stated earlier, the arithmetic mean estimate of the drift \( m^*_a(x) \) presented in Fig. 4 shows an isotropic distribution relative to the distance from the gravity center, justifying a 1D modeling approach.

To estimate the drift \( m^*_k(r) \), a kernel of bandwidth 35 m was selected to compute tuna density estimates in distance classes of 5 m width (Fig. 5). Values of bandwidth and distance class width were adjusted to select the best-fitting model (highest R-squared) of the kernel-estimated drift with the advection-diffusion model. The advection-diffusion ratio \( ADR (=D/|u|) \) varied with the distance from the aggregation gravity center (Fig. 5). The ratio decreased in a central region within 120 m from the aggregation center and showed some low amplitude oscillations beyond 120 m. The \( ADR \) decreased sharply within 30 m from the center and more steadily between 30 to 120 m. Within the model framework, the diffusive component (random individual tuna trajectories) would therefore be higher close to the aggregation center and would rapidly decrease toward the edges. Fluctuations in gradients usually materialize boundaries of system elements (Allen and Starr 1982). \( ADR \) oscillations observed beyond 120 m from the aggregation center could materialize the aggregation boundary i.e. the limits of the area of applicability of the advection-diffusion model.

The \( ADR \) displayed a strong relationship with the drift of the tuna density (Fig. 6). The relationship was modeled on a log scale using two linear models \( lm_1 \) and \( lm_2 \), fitted over two distinct parts of the curve (Fig. 6). The analytical model characterising the relationship between \( ADR \) and \( m^*_k(r) \) was
\[
\begin{align*}
\log(D/|u|) &= \alpha_1 \log(m^*_k) + \beta_1 \Leftrightarrow D/|u| = (m^*_k)^{\alpha_1} e^{\beta_1} \quad \text{for} \quad m^*_k \in [m^*_{k,\min};m^*_{k,\max}] \\
\log(D/|u|) &= \alpha_2 \log(m^*_k) + \beta_2 \Leftrightarrow D/|u| = (m^*_k)^{\alpha_2} e^{\beta_2} \quad \text{for} \quad m^*_k \in [m^*_{k,\min};m^*_{k,\max}],
\end{align*}
\]

where \( m^*_{k,\min} \) and \( m^*_{k,\max} \) are the minimum and maximum \( m^*_k \) values, and \( m^*_{k,L} \) is the value of \( m^*_k \) at the intersect between \( lm_1 \) and \( lm_2 \).

By replacing \( ADR \) in Equation 7 by its expression in Equation 21 and after integration, we obtain an analytical model \( m^*_{\text{mod}} \) for the drift

\[
\begin{align*}
m^*_{\text{mod}} &= \left((m^*_{k,\max})^{\alpha_1} - \alpha_1 r \times e^{-\beta_1}\right)^{1/\alpha_1} \quad \text{for} \quad r \in [r_{\min};r_L] \\
m^*_{\text{mod}} &= \left((m^*_{k,L})^{\alpha_2} - \alpha_2 (r - m^*_{k,L}) \times e^{-\beta_2}\right)^{1/\alpha_2} \quad \text{for} \quad r \in [r_L;r_{\max}],
\end{align*}
\]

where \( r_{\min} \) and \( r_{\max} \) are the minimum and maximum distances from the gravity center of the aggregation, while \( r_L \) is the distance for which \( m^*_k = m^*_{k,L} \).

The intercept between \( lm_1 \) and \( lm_2 \) was defined to minimize the residual sum of squares of the fit of \( m^*_k(r) \) with \( m^*_{\text{mod}}(r) \). The best fit provided a \( R^2 \) value of 0.99 and was obtained at \( r_L = 45 \) m (Fig. 7).

No relation was found between the variance of residuals \( R^*_m(x,t) \) computed in time and the drift value at point \( x \). The universal kriging model was therefore additive. We computed means weighted by area of influences for each map of residuals. The average value of map means was close to zero (0.12, SD = 0.12), meaning that the advection-diffusion model estimate of the drift was unbiased.
336 Spatial correlation modeling

337 Fig. 8 shows the residuals variogram with its fit. The model was fit using the goodness of fit criterion
338 proposed by Rivoirard et al. (2000). The fit model \( \gamma \) is spherical with a nugget of 2, a range of 52 m and
339 the sill at 1.8. The spherical component of the variogram accounts for 50% of the total modeled
340 variance (i.e. the sums of sills). This indicates that the residuals are significantly spatially correlated,
341 justifying the universal kriging approach.

342 The average empirical residual variance \( \bar{s}_R^2(V) \) (3.87) compared well with its model-based estimate
343 \( \bar{\gamma}_{vV} \) (3.73), meaning that the model residual variance is on average well scaled over all realizations.
344 Considering variance in particular surveys, the distribution of \( D_R \) was skewed towards large positive
345 values (Fig. 9), corresponding to surveys with relatively high \( s_R^2(t) \) values. However, its mean was
346 close to zero (0.15, SD = 2.5) and represented only 4% of the mean \( s_R^2(t) \) value. The residual variance
347 unexplained by the spatial correlation model \( \gamma \) was therefore generally low. The spatial model was
348 hence considered appropriate to represent the survey data.

349 Precision of daytime star surveys

350 The estimation variance term associated with the mean residuals computed using the variogram model
351 in Equation 17, was \( \sigma_R^2 \). The global estimation variance for any of the daytime star surveys was

\[
\sigma_E^2 = E[Z_v^2(t)] \times \sigma_{spa}^2 = E[Z_v^2(t)] \times \left( \sigma_m^2 + \sigma_R^2 \right) = 3.58 \times 10^{-10} \times (0.01 + 0.02) = 1.07 \times 10^{-11},
\]

352
Estimates of daily maximum abundance

Table 2 presents 11 estimates of tuna daily maximum abundance around the two moored FADs surveyed from April, 2003, to August, 2003. Daily maxima of abundance were generally observed around noon (mean time: 12:27, SD = 03:09) (Table 2). The mean estimate of tuna daily maximum abundance was 9 MT, with no significant differences noted between the two surveyed FADs.

In order to obtain information on the way kriging weights operated on the different surveys, we computed the average of the weights in classes of distance from the head of the FAD. Kriging weights decreased towards the center of the domain \( V \) in all of these surveys. The mean weights taken over all surveys (Fig. 10) illustrate that trend. Kriging weights therefore correct the oversampling of the center of the area.

Sampling effort and star survey precision

Fig. 11 shows the residual estimation variance \( \sigma^2 \) obtained with different star survey designs as a function of sampling effort. Designs star2d and star4, as well as designs star4d and star8 have similar effort but different number of branches. Designs with higher number of branches yielded the lowest estimation variance as they sample space more evenly.

A total of three star designs with duplicate (star4d, star6d and star8d) and two designs without
duplicate (star6 and star8) yielded low estimation variances. The star8 survey design (Fig. 1b) provided a relatively low estimation variance with a moderate sampling effort. It could provide a good trade-off between precision and sampling effort but could be more difficult to complete than the star8d design in strong current conditions.

Discussion

To our knowledge, this study is the first attempt to define a methodology for analyzing acoustic data obtained during star surveys to estimate the abundance of tuna aggregations around moored FADs. Based on repeated surveys, we observed that the tuna aggregations around the two FADs showed a predictable distribution pattern in which fish abundance decreased with the distance from the head of the FAD. The distribution pattern was modeled with a generic advection-diffusion model of animal grouping. We made use of the universal kriging method which allowed dealing with the deterministic spatial component as well as residual spatial variations. The precision of the survey estimate then depended not on all the spatial variability but only on that of the residuals, theoretically giving a lower estimation variance. We considered repeated surveys as different realizations of the aggregation phenomenon around the FADs and used the repetitions to infer an average non-stationary spatial model. Our (average) model can be used to map the aggregation around the FAD and estimate its abundance and estimation variance. We used the model for an average survey set up and compared different survey designs. We also used the model for particular surveys and performed the estimation for these using the average model and the particular sample locations for these surveys.

Assuming that these two surveyed FADs are representative of others, the procedure for analyzing repeated star surveys around moored FADs can be summarized as follows (Fig. 2): (a) record the
position of the head of the FAD during the survey, (b) center the sample positions relative to the FAD’s head, (c) normalize the sample positions using the standard deviation of positions, (d) grid the survey area and estimate the average fish density in each cell, (e) normalize each cell value by the survey average and estimate relative densities, (f) perform (a) to (e) for each survey of a series of repeated surveys and estimate the average relative density surface across surveys, (g) model the deterministic component in the relative density surface: model the decrease in average relative density from the center of the aggregation to its borders using a 1D advection-diffusion model, (h) estimate 2D residuals in each cell for each survey, (i) estimate a variogram for the residuals, and (j) estimate the survey mean and its precision by kriging. In the previous procedure we used different weightings at different steps. In step (e) we used the simple data average to normalize each cell value, mainly for simplicity's sake. In step (i) we used spatial weights (area of influence) to estimate the residual variogram as suggested in Rivoirard et al. (2000). In step (j) we used kriging weights to estimate the survey mean around the FAD and its precision.

Using simulations, Doonan et al. (2003) tested different survey plans and analysis methods for star surveys. They used a bi-gaussian function to parameterize the decrease in abundance from the center to the border of the aggregation. The advection-diffusion model used here is biologically more informative and also more generic than a bi-gaussian surface. The advection-diffusion model can adapt to a variety of gaussian-like spatial patterns in the data where the shape in the decrease of the abundance from the center to the border is parameterized by the relationship between the $ADR$ (advection diffusion ratio) and the abundance. Doonan et al. (2003) also suggested transforming star survey sample cartesian coordinates to polar coordinates, resulting in transforming the star design into
a series of parallel transects. Though seemingly practical, this procedure will result in producing a discontinuity in the transformed design at the middle of the star design, exactly where the fish aggregation is more continuous. This is thought to result in increasing the estimation variance. Kriging (by weighing the samples optimally) will counterbalance the oversampling effect of the star survey at the middle of the star and we see no need to transform the coordinates to resolve that particular effect.

In our analysis, the most important aspect to deal with was the correct estimation and interpretation of the deterministic component in the aggregation around the head of the FAD. The drift component was successfully modeled within an advection-diffusion framework using 50 star surveys conducted over a four month period. How many surveys are indeed required to infer the drift? To answer this question, we applied the mean precision estimation methodology to subsets of 40, 30, 20, 10 and 5 surveys randomly selected between the initial 50 surveys to estimate the amount of surveys required to get a realistic drift model, as well as a reasonable level of precision, The drift was successfully modeled with ten surveys and more, using kernel bandwidths ranging from 35 to 40 m. The mean surface density computed with five surveys showed too much anisotropy to allow for the fitting of the advection-diffusion model, whatever the bandwidth used in the kernel estimation procedure. Variogram models of residuals and precision levels were similar while using 20 surveys and more. The spherical component of the variogram model computed with 10 surveys accounted for only 33% of the sums of sills. This absence of strong spatial structure in the residuals makes the use of the universal kriging model questionable in this instance. The proportion of total variance assigned to spatial autocorrelation in the variogram model can hence provide guidance to decide whether to use the universal kriging model. These results suggest that our methodology requires a minimum of 20 surveys to provide a reasonable
estimate of the drift and good precision levels. However, over larger time and spatial scales, characteristics of the spatial aggregation can be expected to vary (e.g., seasonal and or regional variations in the drift as well as in the residual spatial correlation). We therefore suggest using the distribution of $D_R(t)$ as a quality control criterion to judge for the adequacy of the average spatial model for each of the repeat cruises. In the larger ensemble of cruises around the two FADs in Martinique, data from 5 surveys were not well fitted with the model presented in this study. They showed exceptionally dense and localized abundance patches and were characterized by high residual variance with no spatial correlation in the residuals. Their geostatistical spatial aggregation index (Petitgas 1998) was significantly higher than that of other surveys (see test in Petitgas 1998). As the fish density surface of these surveys dramatically departed from the drift, they were excluded from the present analysis and would require a separate analysis.

The high regularity of the mean tuna density surface indicates that the density surfaces of tuna aggregations observed during each survey were roughly similar. This suggests that the spatial distribution of tuna observed at the macroscopic scale of the area $V$ during each survey may result from identical, time invariant processes occurring at the microscopic scale of individual tunas. The fact that the drift component in the spatial distribution was successfully modeled within an advection-diffusion framework suggests that these microscopic time invariant processes could be analogous to advection-diffusion. Within the framework of the model, the dome-shaped, tails truncated spatial distribution of midge and tuna aggregations is explained by the dominance of diffusion near the center of the aggregation and by the increase of advective processes that generate sharp edges at the aggregation boundaries. Okubo and Chiang (1974) validated their model by directly estimating realistic advection
and diffusion coefficients based on individual midge velocities. In the same way, studying the small-scale 3D individual trajectories of tuna aggregated around a FAD with acoustic tags should allow validation of the use of the Okubo et al. (2001) advection-diffusion model in the case of tuna aggregations. Moreover, Okubo et al.’s (2001) model is valid for stable groups with no gain and loss of animals. Most studies on residence time of tuna around moored FADs (Musyl et al. 2003; Girard et al. 2004; Ohta and Kakuma 2005) suggest that this hypothesis may hold at the temporal scale of a star survey (30 minutes). However, tuna aggregations around moored FADs are thought to be at least partly maintained at a larger temporal scale by a dynamical equilibrium between emigrating and immigrating fishes (Girard et al. 2004; Doray et al. 2006). Therefore, more realistic models of tuna aggregation processes around FADs should include the large scale gain and loss of individuals.

When large numbers of FADs equipped with radio beacons are deployed, searching time no longer provides a measure of fishing effort, as FADs concentrate fish in known areas. As a consequence, it is now very difficult to use the catch per unit effort as an index of local abundance in stock assessment models (Ariz Telleria et al. 1999; Fréon and Misund 1999; Fonteneau et al. 2000). Stock assessment models that take into account the aggregative behavior of large pelagic fish have been developed (Clark and Mangel 1979; Samples and Sproul 1985; Hilborn and Medley 1989), but their implementation is hindered by the lack of field estimations of exchange rates between aggregated and non-aggregated populations (Fréon and Misund 1999). Our methodology could serve to compute precise abundance estimates around drifting FADs. This would be a first step toward quantitative field assessment of gain and loss of fishes associated with drifting FADs. These flux estimates could be used as proxys of exchange rates between aggregated and non-aggregated populations in stock assessment models. In the
same way, if one assumes that the tuna abundance in the close vicinity of a FAD is proportional to the mesoscale abundance of aggregated tuna in open waters surrounding the device, precise local abundance estimates obtained around FADs could be used to get insights into the distribution and abundance of aggregated tuna in global pelagic ecosystems.

Acknowledgments: This research was co-funded by the European Community, the Regional Council and the prefecture of Martinique. The authors gratefully thank the crew of the “Béryx” for their invaluable assistance during the sea cruises, as well as Eric Mittlestaed and Gene Vricella for correcting the English text. The referees and the editor are thanked for thoughtful comments that greatly improved the manuscript.
References


In Proceedings of the Tuna Fishing and Fish Aggregating Devices Symposium, Trois Ilets, Martinique pp. 15-35.


# Tables

## Table 1. List of the star survey designs tested for their precision.

<table>
<thead>
<tr>
<th>Code</th>
<th>Survey designs (centered around virtual FAD)</th>
<th>Sampling effort (Nb. of sampled cells)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Star2d</td>
<td>2 branches star pattern, with duplicate</td>
<td>128</td>
</tr>
<tr>
<td>Star4d</td>
<td>4 branches star pattern, with duplicate</td>
<td>256</td>
</tr>
<tr>
<td>Star6d</td>
<td>6 branches star pattern, with duplicate</td>
<td>384</td>
</tr>
<tr>
<td>Star8d</td>
<td>8 branches star pattern, with duplicate</td>
<td>512</td>
</tr>
<tr>
<td>Star2</td>
<td>2 branches star pattern, without duplicate</td>
<td>64</td>
</tr>
<tr>
<td>Star4</td>
<td>4 branches star pattern, without duplicate</td>
<td>128</td>
</tr>
<tr>
<td>Star6</td>
<td>6 branches star pattern, without duplicate</td>
<td>192</td>
</tr>
<tr>
<td>Star8</td>
<td>8 branches star pattern, without duplicate</td>
<td>256</td>
</tr>
</tbody>
</table>
Table 2. Estimates of tuna maximum daily abundance around moored FADs.

<table>
<thead>
<tr>
<th>FAD</th>
<th>Date</th>
<th>Time</th>
<th>$Z_a^*$</th>
<th>$Z_v^*$</th>
<th>$\sigma^2$</th>
<th>$\sigma_{spa}^2$</th>
<th>SE</th>
<th>$\rho_a$ (Nb. Fish/m²)</th>
<th>A (Nb. fish)</th>
<th>B (MT)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coastal</td>
<td>29 Apr. 2003</td>
<td>12:30</td>
<td>$3.84 \times 10^{-05}$</td>
<td>$2.41 \times 10^{-05}$</td>
<td>$6.48 \times 10^{-11}$</td>
<td>0.04</td>
<td>33%</td>
<td>0.08</td>
<td>6130</td>
<td>17</td>
</tr>
<tr>
<td>Coastal</td>
<td>3 Jul. 2003</td>
<td>07:45</td>
<td>$1.41 \times 10^{-05}$</td>
<td>$1.34 \times 10^{-05}$</td>
<td>$1.28 \times 10^{-11}$</td>
<td>0.06</td>
<td>27%</td>
<td>0.04</td>
<td>3409</td>
<td>9</td>
</tr>
<tr>
<td>Coastal</td>
<td>4 Jul. 2003</td>
<td>15:30</td>
<td>$9.52 \times 10^{-06}$</td>
<td>$7.58 \times 10^{-06}$</td>
<td>$4.80 \times 10^{-12}$</td>
<td>0.05</td>
<td>29%</td>
<td>0.02</td>
<td>1928</td>
<td>5</td>
</tr>
<tr>
<td>Coastal</td>
<td>11 Jul. 2003</td>
<td>12:20</td>
<td>$2.33 \times 10^{-05}$</td>
<td>$1.62 \times 10^{-05}$</td>
<td>$3.21 \times 10^{-11}$</td>
<td>0.06</td>
<td>35%</td>
<td>0.05</td>
<td>4128</td>
<td>11</td>
</tr>
<tr>
<td>Coastal</td>
<td>3 Aug. 2003</td>
<td>16:30</td>
<td>$1.71 \times 10^{-05}$</td>
<td>$1.61 \times 10^{-05}$</td>
<td>$1.36 \times 10^{-11}$</td>
<td>0.05</td>
<td>23%</td>
<td>0.05</td>
<td>4098</td>
<td>11</td>
</tr>
<tr>
<td>Coastal</td>
<td>4 Aug. 2003</td>
<td>17:00</td>
<td>$2.18 \times 10^{-06}$</td>
<td>$1.17 \times 10^{-06}$</td>
<td>$2.09 \times 10^{-13}$</td>
<td>0.04</td>
<td>39%</td>
<td>0.00</td>
<td>297</td>
<td>1</td>
</tr>
<tr>
<td>Coastal</td>
<td>5 Aug. 2003</td>
<td>12:00</td>
<td>$1.48 \times 10^{-05}$</td>
<td>$1.07 \times 10^{-05}$</td>
<td>$1.01 \times 10^{-11}$</td>
<td>0.05</td>
<td>30%</td>
<td>0.03</td>
<td>2722</td>
<td>7</td>
</tr>
<tr>
<td>Offshore</td>
<td>6 Aug. 2003</td>
<td>12:30</td>
<td>$1.57 \times 10^{-05}$</td>
<td>$1.08 \times 10^{-05}$</td>
<td>$1.02 \times 10^{-11}$</td>
<td>0.04</td>
<td>30%</td>
<td>0.03</td>
<td>2750</td>
<td>7</td>
</tr>
<tr>
<td>Offshore</td>
<td>7 Aug. 2003</td>
<td>07:00</td>
<td>$1.58 \times 10^{-05}$</td>
<td>$1.29 \times 10^{-05}$</td>
<td>$1.14 \times 10^{-11}$</td>
<td>0.05</td>
<td>26%</td>
<td>0.04</td>
<td>3270</td>
<td>9</td>
</tr>
<tr>
<td>Offshore</td>
<td>8 Aug. 2003</td>
<td>11:45</td>
<td>$1.87 \times 10^{-05}$</td>
<td>$1.52 \times 10^{-05}$</td>
<td>$1.59 \times 10^{-11}$</td>
<td>0.05</td>
<td>26%</td>
<td>0.05</td>
<td>3859</td>
<td>10</td>
</tr>
<tr>
<td>Offshore</td>
<td>8 Aug. 2003</td>
<td>12:15</td>
<td>$1.68 \times 10^{-05}$</td>
<td>$1.44 \times 10^{-05}$</td>
<td>$1.23 \times 10^{-11}$</td>
<td>0.04</td>
<td>24%</td>
<td>0.05</td>
<td>3652</td>
<td>10</td>
</tr>
<tr>
<td>Mean</td>
<td></td>
<td>12:27</td>
<td>$1.69 \times 10^{-05}$</td>
<td>$1.30 \times 10^{-05}$</td>
<td>$1.71 \times 10^{-11}$</td>
<td>0.05</td>
<td>29%</td>
<td>0.04</td>
<td>3295</td>
<td>9</td>
</tr>
</tbody>
</table>
Figure captions

Fig. 1. Star survey designs. (a) 8 branches with duplicate (star8d) (b) 8 branches without duplicate (star8).

Fig. 2. Schematic outline of the methodology designed to estimate the mean density and precision of star acoustic surveys of tuna aggregations associated to moored fish aggregating devices.

Fig. 3. Example of acoustic tuna density map obtained during a star survey around a moored fish aggregating device (FAD) in Martinique. The radius of each circle is proportional to $s_a$ values (m$^2$•m$^{-2}$) recorded in each elementary sampling unit.

Fig. 4. Spatial distribution of the arithmetic mean of relative acoustic tuna density recorded during daytime surveys in the longitudinal (solid line) and latitudinal (broken line) directions.

Fig. 5. Kernel (empty circles) and arithmetic (filled circles) one-dimensional estimate of the mean relative tuna density, with the advection-diffusion ratio overlaid (triangles).

Fig. 6. Advection-diffusion ratio as a function of tuna kernel density estimate (empty circles) with the 2 parts-linear model overlaid (broken lines).

Fig. 7. Fit of the advection-diffusion model (solid line) with the one-dimensional distribution of tuna kernel density estimate (empty circles), $R^2 = 99\%$.

Fig. 8. Mean experimental variogram of residuals (dots). The fitted model (solid line) is a nugget of value 2 added to a spherical model of sill 1.8 and range 52 m.

Fig. 9. Boxplot of the differences $D_R(t)$ between the empirical residual variance $s_R^2(t)$ and the model-
based variance $\gamma_{v}$, for each survey $t$. The horizontal bold line represents the median, the lower and upper edges of the box the first and third quartile, and the lower and upper whiskers the limits of the 95% confidence interval for the median. Empty dots are outliers lying beyond the extremes of the whiskers.

**Fig. 10.** Mean krigging weights as a function of the distance from the head of the fish aggregating device (FAD).

**Fig. 11.** Residual estimation variance $\sigma^2_r$ as a function of sampling effort for different star survey designs described in Table 1.
Figures

(a) (b)

Fig. 1.
Fig. 2.

**Single survey**

- Center the sample positions relative to the FAD’s head positions
- Normalize the sample positions
- Grid the survey area and estimate the average fish density in each cell
- Estimate relative densities, define aggregation limits
- Estimate 2D residuals in each cell for each survey
- Estimate the survey mean and its precision by kriging
- Estimate the aggregation biomass using mean individual TS and weight

**N surveys**

- Estimate the average relative density surface across surveys
- Model the decrease in average relative density from the center of the aggregation to its borders using a 1D advection-diffusion model
- Estimate a variogram model for the residuals
- Estimate the mean survey precision by kriging
- Estimate precision of other survey designs
Fig. 3.
Fig. 4.

Distance from gravity center (m)

Mean relative acoustic density

Distance from gravity center (m)
Fig. 5.
Fig. 6.
Fig. 7.
Fig. 8.
Fig. 9.
Fig. 10.
Fig. 11.

The figure shows a scatter plot with the number of sampled cells on the x-axis and some unknown value on the y-axis (labeled as $\gamma$). The plot includes different symbols for different categories: star2d, star4d, star6d, star8d, star2, star4, star6, and star8. The trend lines suggest a positive correlation between the number of sampled cells and the value on the y-axis.