



**HAL**  
open science

# Unsupervised Bayesian reconstruction of individual life histories from otolith signatures : case study of Sr : Ca transects of eel (*Anguilla anguilla*) otoliths

Ronan Fablet, Françoise Daverat, Hélène de Pontual

► **To cite this version:**

Ronan Fablet, Françoise Daverat, Hélène de Pontual. Unsupervised Bayesian reconstruction of individual life histories from otolith signatures : case study of Sr : Ca transects of eel (*Anguilla anguilla*) otoliths. *Canadian Journal of Fisheries and Aquatic Sciences*, 2007, 64 (1), pp.152 - 165. hal-02344299

**HAL Id: hal-02344299**

**<https://hal.science/hal-02344299v1>**

Submitted on 4 Nov 2019

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Unsupervised Bayesian reconstruction of individual life histories from otolith signatures: case study of Sr:Ca transects of eel (*Anguilla anguilla*) otoliths.

5 **Ronan Fablet<sup>1\*</sup>, Françoise Daverat<sup>2</sup> and Hélène De Pontual<sup>1</sup>.**

1. Ifremer/LASAA, BP 70, 29280 Plouzane, France

2. Cemagref, 50 Avenue de Verdun, 33612 Cestas, France

Abstract: The reconstruction of individual life histories from otolith measures was stated as  
10 an unsupervised signal processing issue embedded in a Bayesian framework. The proposed  
computational methodology was applied to a set of 192 eel (*Anguilla anguilla*) otoliths. It  
provided an objective, robust and unsupervised analysis of the whole set of individual  
chronologies of habitat use (either river, estuary or coastal) from chemical otolith signatures,  
given as Sr:Ca measures acquired along an otolith growth axis. To this end, links between  
15 Sr:Ca values and habitat, age and season, as well as the likelihood of the transitions from one  
habitat type to another were modelled. Major movement characteristics such as age at  
transition between habitats and time spent in one habitat were estimated. As a straightforward  
output, an unsupervised classification of habitat use patterns was determined and showed a  
great diversity. A total of 37 patterns of habitat use were found among which 20 different  
20 patterns accounted for 90% of the sample. In accordance with results from literature,  
residence behaviour for each habitat type was observed (28% of the samples). However, about  
72% changed habitat once or of several times, mainly before age four. This method opens  
new avenues for the analysis of individual habitat strategies. In addition, it could be easily  
applied to any other measures (microchemistry or not) taken along an otolith growth axis to  
25 reconstruct life histories.

KEY WORDS: individual life histories, fish otoliths, Bayesian labeling, Gaussian mixture models, hidden Markov models, otolith microchemistry.

---

\* Email : [rfablet@ifremer.fr](mailto:rfablet@ifremer.fr)

## Introduction

The recent bloom of ecology studies using otolith microchemistry emphasises the remarkable potential of the otolith accuracy for investigating fish life history traits. The use of otolith chemistry to trace migration pathways is premised on a significant correlation between the elemental composition of otoliths and physicochemical properties of the ambient environment (Campana 1999; Martin and Thorrold 2005; Thorrold et al. 1997). Hence, sequential elemental measures acquired along an otolith growth axis are thought to record environmental information along the fish lifespan. River basin origin have been inferred from strontium isotopes ratios (Milton and Chenery 2003; Thorrold and Shuttleworth 2000), while salinity level has been inferred from ratios of strontium on calcium (Tzeng et al. 1994). With more than 200 research articles published, strontium calcium ratios became the most popular microchemistry application to fish ecology, as a tool to track movements across water masses of different salinity for at least 20 fish species. To date, more than 28 published studies have exploited strontium calcium ratios for the analysis of life histories of eels species (*Anguilla spp*). Data acquisition of otolith microchemistry remains technically, timely and costly demanding. Hence, the challenge has first been a matter of data acquisition rather than a matter of data interpretation. So far, most otolith microchemistry transects of measures were interpreted by a visual evaluation of each individual signal. Whenever formalised, the treatment of sequential Sr:Ca measures of eel otoliths data consisted in calculation of the mean values of each individual eel (Tsukamoto and Arai 2001, Tzeng et al. 2002). For instance, Tzeng et al. 2002 classified eel life histories from the number of years assigned to a specific water mass according to the mean value of Sr:Ca measures. This is questionable as each Sr:Ca otolith measure is a specific indicator of a water mass so that the mean of two different water masses has no ecological meaning. Besides, the temporal dimension of the data was lost. Other studies interpreted directly individual Sr:Ca transects values plotted

against age class graphs, with Sr:Ca values assigned to a water mass according to their level (Morrison et al. 2003). This was not either satisfactory, as, due to the non-linearity of the otolith growth pattern, the first years of the fish life have a greater spatial resolution than the last years of the fish life. As a consequence, evenly spaced Sr:Ca measures transects result in  
5 non equal number of measures for each year of the fish life. In addition, the classification of individual migratory behaviours relied on *a priori* classes, which may not account for the actual diversity observed in the processed dataset ( Tzeng et al. 2002, Daverat and Tomas *in press*). For instance, a large data set of 270 eel Sr:Ca transects was classified with a supervised method accounting for the mean and the variations of each individual curve  
10 (Daverat and Tomas *in press*). From classes of life histories defined *a priori*, the classification of the eels resulted from a visual interpretation of the associated Sr:Ca signal. While time consuming, such a scheme also appears rather subjective.

It emerged from this overview of previous work a need for an unsupervised and well-founded computational method that could at the same time account for the temporal nature of the  
15 sequential measures and infer a relevant interpretation of the chemical signal (Sr:Ca ratio) in terms of environmental information (here the habitat visited by the fish). Signal processing methods appeared as a promising tool to tackle this issue. Multidisciplinary studies have already applied signal processing techniques to process fish otolith data for ageing purposes (Troadek et al. 2000; Fablet 2005) or other issues such as stock discrimination (Campana and  
20 Casselman 1993), or fish individual status (Cardinale et al. 2004). Within a signal processing framework, each Sr:Ca measure is associated with a hidden state variable standing for an environmental information (in our case, an habitat and the associated water mass), and the temporal nature of the sequence of Sr:Ca measures can be restored from the otolith growth pattern. Formally, the reconstruction of the individual patterns of habitat use is stated as the  
25 Bayesian reconstruction of the temporal sequence of the hidden state variables from the

observed sequence of Sr:Ca signatures. The proposed scheme mainly relies on Gaussian mixture models and hidden Markov models. All these developments were implemented under Matlab 7 using Netlab (Nabney 2001) and CRF (Murphy 2004) toolboxes. This method was applied to the interpretation of 192 eel otolith Sr:Ca transects of measures in terms of individual habitat use histories. The generalisation of such a method for the treatment of sequential measures taken on an otolith growth axis is further discussed.

## **Material and Methods**

### **Data sets**

An eel ecology study (Daverat et al. 2005), led to the acquisition of 192 individual eel Sr:Ca series. The fish samples were collected in the Gironde river basin SW France in three main habitats (water masses). 63 eels (33%) were collected in freshwater habitats, 115 eels (60%) in the estuary habitats and 14 eels (7%) in the coastal habitats. The aim of this study was to characterize the habitat use patterns of eels from the Gironde river basin, during their continental growing phase as a yellow eel. Sr:Ca ratios transects were used to track the eels movements across freshwater, brackish and marine habitats. Hence, only the part of the otolith corresponding to the continental life of the eel was retained (from glass eel mark to the edge). The acquisition method was described in Daverat et al. (2005), and consisted in electron microprobe measures of Sr and Ca concentrations in 8  $\mu\text{m}$  diameter spots, evenly spaced every 20  $\mu\text{m}$  along the otolith longest growth axis from the glass eel mark to the edge of the otolith. Along this transect of Sr:Ca measures, the position of each annual age mark was recorded as a distance from the glass eel mark.

### **Calibration over time of Sr :Ca series .**

The interpretation of the macrostructures, so-called “rings”, laid annually (Berg 1985) observed on the otoliths along the longest growth axis, provided an estimation of individual growth patterns. The elver mark was set as the origin of the time axis and only the interval

between the elver mark and the edge was taken into account, as the ecological issue was the continental habitat use pattern of the eel after the glass eel stage until the time of capture. Annual rings were used as time references to transform Sr :Ca series acquired with respect to the distance to the elver mark to time series using a linear interpolation. The time series was  
5 interpolated at a monthly precision (that is to say a time sampling rate of 1/12). In the following, we will refer to this time axis as the age axis, since it refers to the time spent from the elver mark.

The actual temporal resolution of the Sr:Ca series depends both on the sampling resolution of the electron microprobe and obviously of the otolith growth rate. In a previous experiment  
10 (Daverat et al. 2005), a mean otolith growth rate of 20  $\mu\text{m}$  per month was found for young individuals, so that about 11 Sr:Ca measures are usually sampled for the first growth years, whereas from the 6<sup>th</sup> year fewer measures (down to 3 or 4) may be available due to slower growth. Therefore, the results issued from the analysis of interpolated Sr:Ca time series need to be cautiously analyzed in terms of temporal precision, especially for the last years of the  
15 life of the older individuals.

### **Determination of habitat-related Sr:Ca model**

Following Daverat et al. (2005), Sr :Ca measures can be regarded as a proxy of the habitat type for eels of the Gironde watershed. Three habitat categories were considered according to  
20 salinity compartment: river, estuary and marine habitats. We further modelled the distribution of Sr:Ca signatures for each habitat as a Gaussian distribution parameterized by a mean model and a standard deviation.

Different models could be chosen. In this study, two different cases were investigated. The first one was a constant model parameterized by a mean value. In order to test for the  
25 influence of seasonality and age on the incorporation of strontium, a linear model with two

explanatory variables (fish age and month) was also considered. Formally, let us denote by  $g(\cdot|\Theta_H, \sigma_H)$  the Gaussian distribution of Sr:Ca measures for habitat H, parameterized by the mean model  $\Theta_H$  and the standard deviation  $\sigma_H$ . Using a constant model  $\Theta_H = m_H$ ,  $g(\cdot|m_H, \sigma_H)$  is computed for a Sr:Ca measure  $y$  as:

$$5 \quad g(y|m_H, \sigma_H) = \frac{1}{\sqrt{2\pi\sigma_H^2}} \exp\left(-\frac{(y-m_H)^2}{2\sigma_H^2}\right)$$

Considering a linear model,  $\Theta_H$  is explicitly defined by the mean value  $m_H$ , the effect of the age  $\lambda_A$  and the effect of the season  $\lambda_S$ . For a Sr:Ca measure  $y$  at age  $a$  and hydrological season  $s$  (normalized average monthly flow), the associated likelihood  $g(y|a, s, \Theta_H, \sigma_H)$  is given by:

$$g(y|a, s, \Theta_H, \sigma_H) = \frac{1}{\sqrt{2\pi\sigma_H^2}} \exp\left(-\frac{(y-m_H-\lambda_A a-\lambda_S s)^2}{2\sigma_H^2}\right)$$

10 The constant model is a particular case of the linear model with  $\lambda_A = \lambda_S = 0$ . Hence, in the subsequent, we will only detail the developments for the latter.

As a first step, we aim at determining for each habitat type H the associated model parameters  $(\Theta_H, \sigma_H)$ . One should note that no labelled data is available to perform this estimation, but only a mixed set of Sr:Ca measures  $\{y_i\}$  associated with unknown habitats (within the three  
15 considered ones). The estimation of the parameters of the habitat models  $\{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}}$  is then stated as an unsupervised issue. To this end, given the Sr:Ca measures  $\{y_i\}$  relative to explanatory variables  $\{a_i, s_i\}$ , the whole distribution of  $\{y_i\}$  is modelled as a Gaussian mixture issued from the superimposition of the three Gaussian habitat models:

$$20 \quad p(y_i|a_i, s_i, \{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}}) = \sum_{H \in \{R, E, M\}} \pi_H g(y_i|a_i, s_i, \Theta_H, \sigma_H),$$

where R, E and M stand for the labels relative to the three habitats: respectively, river (R), estuary (E) and marine area (M).  $\pi_R, \pi_E, \pi_M$  are the prior probabilities for each habitat. Given

$\{y_i\}$  and  $\{a_i, s_i\}$ , we aim at estimating the parameters of the mixture model  $\{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}}$  such that  $p(y|a, s, \{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}})$  best fits to the distribution of the dataset  $\{y_i\}$  according to the maximum likelihood (ML). This model estimation is carried out using the EM (Expectation-Maximization) algorithm (Bishop 1995). The computations involved in this iterative procedure are detailed in Annex I.

The estimated mixture parameters are finally assigned to each habitat by sorting the mean values of the Gaussian modes. We rely on the statement that the lower the salinity of the habitat the lower the mean Sr:Ca measure (Fig. 1).

The goodness of the fits of the constant model and the linear model for each mode was compared with AIC (Akaike Information Criterion) method (Awad 1996) and the influence of age and season on Sr:Ca value was tested according to correlation statistics between the model prediction and the data (McCullagh and Nelder 1989). The comparison of the contributions of each group of predictors permitted to evaluate the relative importance of habitat, season and age (Silber et al. 1995). These statistical tests were performed with R software (RDevelopmentCoreTeam 2005).

### **Estimation of individual habitat use from Sr:Ca series**

Our goal was to analyze the individual patterns of habitat use from the series of Sr:Ca measures as illustrated (Fig.2). This issue resorts to the estimation of the temporal sequence of the habitat-related path and is formally stated as a Bayesian labelling issue, that is to say retrieving the temporal habitat sequence  $\{x_t\}$  corresponding to a given observed series of Sr:Ca measures  $\{y_t\}$ , where, for each time  $t$ ,  $x_t$  is a label: R (river), E (estuary) or M (Marine area).

Within a Bayesian framework, this labelling issue comes to the determination of the best sequence  $(\hat{x}_0, \dots, \hat{x}_T)$  according to the Maximum A Posteriori (MAP) criterion, that is to say



retrieving the temporal sequence of habitat categories corresponding to the maximum posterior likelihood given the acquired series of Sr:Ca measures:

$$(\hat{x}_0, \dots, \hat{x}_T) = \arg \min_{(x_0, \dots, x_T)} p(x_0, \dots, x_T | y_0, \dots, y_T)$$

**Eq. 1**

5 Further assumptions are required to solve for this minimization issue. Two solutions are investigated. First, assuming that Sr:Ca measures  $\{y_t\}$  and labels  $\{x_t\}$  are statistically independent. Equation (2) reduces for each time  $t$  to:

$$\hat{x}_t = \arg \min_{H \in \{R, E, M\}} p(x_t = H | y_t)$$

**Eq. 2**

10 Using the estimated Gaussian mixture model  $\{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}}$ , the posterior likelihood function  $p(x_t = H | y_t)$  is computed as:

$$p(x_t = H_1 | y_t, a_t, s_t, \{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}}) = \frac{\pi_{H_1} p(y_t | a_t, s_t, \Theta_{H_1}, \sigma_{H_1})}{\sum_{H_2 \in \{R, E, M\}} \pi_{H_2} p(y_t | a_t, s_t, \Theta_{H_2}, \sigma_{H_2})}$$

This first model is however rather simplistic and does not explicitly model fish movements among habitats. To account for these temporal dynamics, first-order Gaussian hidden Markov  
 15 models (Rabiner 1989) are used. These models were initially developed and exploited for speech analysis. As illustrated (Fig. 3), two main components are involved: a prior on the temporal dynamics of the state variables  $\{x_t\}$  which models fish movements from  $x_{t-1}$  to  $x_t$  and a data-driven term characterizing the probabilistic distribution of the observed Sr:Ca measures  $y_t$  given the habitat type  $x_t$ .

20 The temporal prior is stated as a first-order Markov chain. This resorts to the assumption that, given the sequence of state variables  $(x_0, x_1, \dots, x_{t-1})$  from time 0 to time  $t-1$ , the state variable at time  $t$  only depends on  $x_{t-1}$ . It means that this model only keeps the memory of its last state to jump to the next one. Formally, this leads to the property that

$p(x_t | x_{t-1}, \dots, x_1, x_0) = p(x_t | x_{t-1})$ . Consequently, a first-order Markov chain is fully characterized by its transition matrix  $\Gamma$ :

$$p(x_t = H_1 | x_{t-1} = H_2) = \Gamma(H_1, H_2),$$

which specifies the likelihood that the fish is in habitat type  $H_1$  at time  $t$  given that it is in habitat type  $H_2$  at time  $t-1$ . The graphical representation of the transition matrix is provided (Fig. 4). Let us stress that some transitions may be forbidden, that is to say pair of habitats for which  $\Gamma(H_1, H_2) = 0$  (for instance, the transition from A to C in Fig. 4). However, this does not prevent from reaching one state from another by going through other states, if there is a sequence involving several transitions with a non-null likelihood. For instance as illustrated (Fig. 4), whereas direct transitions from A to C are impossible, paths from A to C going through B are possible.

In addition to the prior component, the data-driven model actually specifies the computation of the likelihood  $p(y_t | x_t)$  of a given measure given its state value. This model resorts to the characterization of the probabilistic distribution of the observed measures for each state. From the Gaussian mixture models of Sr:Ca measures, likelihood  $p(y_t | x_t)$  is formally defined as:

$$p(y_t | x_t = H, \Phi) = g(y_t | a, s, \Theta_H, \sigma_H),$$

where denotes the set of habitat-related Sr:Ca models  $\Phi = \{\Theta_{H_i}, \sigma_{H_i}\}_{H_i \in \{R, E, M\}}$ .

To exploit this Gaussian hidden Markov model, we first need to estimate the parameters of the prior term. Similarly to the estimation of the parameters of the Gaussian mixture model, this estimation is performed according to the ML criterion which resorts to determine the transition matrix associated with the maximum likelihood for the whole set of samples:

$$\hat{\Gamma} = \arg \max_{\Gamma} \prod_i p(x_0^i, \dots, x_T^i | y_0^i, \dots, y_T^i, \Gamma)$$

This maximization issue is solved for using the EM algorithm. We let the reader refer to Rabiner (1989) for a detailed description of this estimation procedure. We only review its

main characteristics. This procedure iterates two steps until convergence. Given the current estimate  $\Gamma^k$  of the transition matrix, the Expectation step resorts to the computation of the posterior likelihoods:

$$\xi_{H_1 H_2}^i(t) = p(x_t^i = H_2, x_{t-1}^i = H_1 | y_0^i, \dots, y_T^i, \Gamma^k),$$

$$\gamma_{H_1}^i(t) = p(x_t^i = H_1 | y_0^i, \dots, y_T^i, \Gamma^k).$$

The M-step follows to update the transition matrix  $\Gamma^{k+1}$  from these posterior likelihoods as their average over the whole dataset:

$$\Gamma(H_1, H_2) = \frac{\sum_{i=1}^N \sum_{t=0}^{T_i} \xi_{H_1 H_2}^i(t)}{\sum_{i=1}^N \sum_{t=0}^{T_i} \gamma_{H_1}^i(t)}.$$

Given the estimate of the parameters of the hidden Markov model, the determination of the optimal MAP sequence  $(\hat{x}_0, \dots, \hat{x}_T)$  defined by  $\mathbb{E}_{\text{q. 2}}$  is solved exactly by the Viterbi algorithm (Rabiner 1989). This algorithm relies on the fact that any subsequence of the optimal MAP sequence is the optimal habitat sequence for the corresponding sequence of Sr:Ca measures. It involves computations similar to the forward procedure of the EM scheme. More precisely, it first computes recursively (from time 0 to time T) the likelihood, denoted by  $\delta_H(t)$ , of the most likely habitat sequence leading to habitat H at time t:

$$\delta_H(t) = \max_{x_0, \dots, x_{t-1}} p(y_0, \dots, y_t, x_0, \dots, x_{t-1}, x_t = H | \Phi, \Gamma).$$

At the final step of the forward procedure (i.e., time T), the habitat label maximizing  $\delta_H(T)$  provides the optimal habitat  $\hat{x}_T$ . A backward procedure then reconstructs recursively the optimal habitat sequence  $(\hat{x}_0, \dots, \hat{x}_T)$  by retrieving the state  $H_1$  at time t-1 which leads to the reconstructed state  $\hat{x}_t$  at time t with the maximum likelihood. We let the reader refer to

Rabiner (1989) for a detailed description of the Viterbi algorithm and of the associated computations.

### **Analysis of habitat sequences**

5 Given the set of the individual habitat sequence, a quantitative analysis of fish movements is carried out. First the global analysis of the habitat sequences delivers a quantitative and unsupervised categorization of the observed movement patterns. The movement pattern is defined as the sequence of the successive habitats visited by the fish. This sequence is defined by the quality and the order of the visited habitats: for instance, the movement pattern issued  
 10 from habitat sequence RRRREEEEERRRRR is RER. Given the whole otolith set, the automated and unsupervised classification of individual movement behaviours can be determined, as well as the relative frequencies of these categories of movement patterns.

Since the habitat sequences are calibrated over time, a variety of measures can also be defined to characterize individual life traits. We focus on the analysis of the time at which the  
 15 transitions from one habitat to another occur, and of the time spent in a given habitat between two transitions. For a given type of transition from habitat type  $H_1$  to habitat type  $H_2$  (i.e., within the set of transitions {R to E, R to M, E to M, E to R, M to R, M to E}), the whole set of habitat sequences  $\{(x_t^i)_{t \in \{0, \dots, T_i\}}\}_{i \in \{1, \dots, N\}}$  is analyzed to extract the set of all the transitions from

$H_1$  to  $H_2$ . These transitions are characterized by their transition times  $\{t_{H_1, H_2}^n\}$  and the times  
 20 spent in  $H_2$   $\{D_{H_1, H_2}^n\}$ . The statistical distributions of these quantities are then computed using a non-parametric technique since they are clearly multimodal. More precisely, given a scale parameter  $\mu$  and a set of quantities  $\{w^n\}$ , the likelihood  $p(w)$  is computed as:

$$p(w) = \frac{1}{Z} \sum_n \exp(-\mu(w - w^n)^2),$$

where  $Z$  is the normalization factor. The computation of the

statistics of transition times  $\{t_{H_1, H_2}^n\}$  is performed in terms of age, month and age group at

which the transitions occur. Scale parameter  $\mu$  is set to 1 for quantities given as monthly values and age groups and to  $1/10^2$  for quantities given as ages.

## **Results**

### **Comparison between constant and linear models.**

- 5 The linear model accounting for habitat, age and season effects (AIC value of 45167) has a better performance than the model only accounting for habitat effect (AIC value of 46465). Habitat, age and hydrological season (river flow) factors all had a significant influence on Sr/Ca values for each habitat ( $p < 0.001$ ) (Table 1).

The comparison of the relative contribution of the habitat factor with the contribution of both  
10 age and season factors revealed that habitat contributes more to the variation of SrCa values than age and season with a ratio of effect standard deviations of 4.58 [95%CI 4.38-4.79], (Silber et al. 1995). Since the effect of age and season is significant in terms p-statistics, we report the results of the analysis of the individual chronologies of habitat with respect to the linear model.

15

### **Individual chronologies of habitat use.**

The comparison of the habitat sequences issued from the Gaussian mixture model and the hidden Markov models showed the improvements brought by the latter one (Fig. 5). As the Gaussian mixture models do not account for time coherence, the associated habitat sequences  
20 may be chaotic and may involve numerous short and unlikely transitions. On the contrary, the explicit modelling of fish movements between habitats thanks to the estimated transition matrix leads to the smoother results reported for the hidden Markov model.

Consequently, this model was chosen to further characterize life traits from the estimated habitat sequences. Eels sampled in the Gironde watershed displayed a wide repertoire of  
25 habitat use patterns, such as residences all life long in the same habitat as defined in the model (river, estuary, marine area) or as single or multiple shifts among habitats (Fig. 6).

### **Analysis and classification of movement patterns**

From the overall analysis of the reduced habitat sequences, the different classes of movement patterns associated with the considered dataset were automatically determined. Our analysis is restricted to fish older than four years and take into account the movement patterns up to age 6. As illustrated, 37 patterns are represented (Fig. 7). While only the first 20 patterns account for more than 80% of the samples, the first five patterns occur with a frequency greater than 5%. Among the first four patterns, three correspond to resident behaviour. This resident behaviour however account for only 28% of the samples and 72% of the movement patterns involve at least one movement. Residents in the marine habitat account for 6.5% of the sample which is quite consistent with the 7% of fish collected in the marine area. While 33% of the sample was collected in the river, only 14.3% of the samples are labeled as residents in the river habitat. Similarly, the reported categorization leads to only 7.8% of residents in the estuary area while our sample is composed of 60% of fish collected in the estuary. Concerning migration behaviours, patterns involving only one migration as well as patterns involving several migrations between two types of habitats are encountered, but only between the river and the estuarine area or between the estuarine and marine areas. Noticeably, up to 5 successive migrations between the river and the estuarine area can be observed for the same individual before age 6. Fewer samples involve migrations among the three habitats. However, about 9% of the samples are associated to patterns including a first migration from the river to the estuarine area, and then one or several movements between the estuarine and the marine areas. Conversely, very few individuals (below 2%) move from the marine area to the river after a stay in the estuarine area.

### **Analysis of the transitions between habitats.**

In order to keep a consistent temporal resolution through the fish life, only the first 6 years of life of the fish were further considered for the rest of the analysis. The proportions of the transitions from one habitat to another one were evaluated first as a function of age and then as a function of season. In the overall sample (192 eels), transitions as a function of age or as a function of season, between two different habitats, were less frequent than stays in the same habitat as presented for instance for the estuary (Fig. 8) at the individual level. This indicates that movements among habitats are seldom compared to the residence in one habitat. Transitions between the river and the marine area did not occur. Besides, the reported results also shows that the occurrence of the transition is season-dependent.

### **Analysis of transitions schedules and duration of habitat use**

Age at transition between the river and the estuary and duration of the habitat use before changing habitat were investigated for eels of more than four years old. Due to their absence, transitions between the marine habitat and the river were not analysed.

Transitions from the river to the estuary did not seem age specific (Fig. 9), but their frequency decreased with age with a maximum of transitions occurring at age 1. The duration of the stay in the river was less than two years with a maximum of fish spending less than one year in the river before moving to the estuary (Fig. 9).

Transitions from the estuary to the river and transitions from the estuary to the marine area did not occur for a specific age of the eel, as shown (Fig. 10) but decreased as the age of the fish increased. Most eels spent less than one year in the estuary before moving either to the river or to the marine area (Fig. 10).

## **Discussion**

### **Spatial and temporal resolution of the analysis**

5 The integration of the temporal dimension in Sr:Ca series was an important issue in this study. Time series were reconstructed using resampling techniques. The spatial resolution of the analysis on the otolith was constrained by analytic requirements. In this study, a measure of 8µm size every 20 µm provided an approximate temporal resolution of one month for the first years of life up three months and more later than the 6<sup>th</sup> year. This constraint leads us to  
10 analyse only the first 6 years of life in order to keep a consistent temporal resolution. Our method was based on a constant growth of the otolith through the year. It would be improved by accounting for the seasonal variations of the eel otolith growth. Despite the general variations of the eel otolith are known (Mounaix and Fontenelle 1994), a formalised model is, as far as we know, not available.

15

### **Habitat-related modelling of otolith signatures**

In this study, a mono-proxy approach was used to model habitat-related otolith characteristics. The proposed unsupervised scheme based on Gaussian mixture models permits to estimate model parameters for each habitat zone from unlabelled data. A linear model was used to  
20 account for the significant influence on Sr:Ca values of age and season in addition to habitat. As expected, the contribution of habitat has been shown to be much greater than that of age and season. Outside the metamorphosis from leptocephalus larvae into glass eel, no significant effect of ontogeny due to growth or age was observed so far on Sr:Ca incorporation into eel otolith (Daverat et al. 2005; Kawakami et al. 1998; Kraus and Secor  
25 2003; Tzeng 1996). A validation using another fish species reared for two years in constant salinity failed in detecting any age effect on Sr:Ca incorporation into otoliths (Elsdon and



Gillanders 2005). Age may affect Sr incorporation at a greater time scale than a few years, especially for some eels that can spend up to 20 years in their feeding habitats. As the mean age of our sample was 7 years, and only the six first years of life were considered, the age effect was weak compared to the habitat effect. The season effect, although very weak, was explained by the seasonal variations of freshwater flows into marine water masses, of the Gironde watershed and by the variations of water temperature. Hence, seasonal variations of the river flow were introduced in the model developed here. In this study site, measures of Sr:Ca ratio in the water, collected at different seasons showed that values were slightly fluctuating over the seasons without affecting the discrimination of marine, estuarine and river habitats (Daverat et al. 2005). As expected, the relation between habitat and Sr:Ca value was very strong, a result validated for eels and other species using coupled experimental and field validations (Daverat et al. 2005; Elsdon and Gillanders 2005; Kraus and Secor 2004). Further applications of this model could take into account other types of effects such as physiology parameters, feeding conditions or temperature (Campana 1999; Campana et al. 2000). In addition, other kinds of parametric models (polynomial, log-normal, exponential, ...) could also be straightforwardly used. Besides, multi proxy approaches using multidimensional structural and/or chemistry otolith signatures may also be investigated within this framework with a view to retrieving a more precise estimation of the environments or states experienced throughout the fish life. The combination of Sr:Ca ratio with oxygen isotopes ratios as a proxy of water temperature (Nelson et al. 1989) may for instance resort to a more precise estimation of the temporal resolution of the measures in the example developed here.

### **Reconstruction of habitats use chronologies**

The proposed method turned out to be particularly adapted to the analysis of large data sets. Our original data consisted in 192 individual Sr:Ca series containing 70 points of Sr:Ca measures on average. Hence a total of 14649 Sr:Ca measures, were analysed as 14649 events

representing an habitat use. The proposed approach is computationally efficient, since only a few minutes are required to process the whole sample set, including both the estimation of the habitat Sr:Ca Gaussian mixture model and the reconstruction of all the individual patterns of habitat use. Compared to previous work, the key feature of this quantitative framework is that a non-subjective interpretation of individual Sr:Ca series in terms of habitat use is provided from an unsupervised analysis.

From a methodological point of view, hidden Markov models have been shown to be much more efficient than Gaussian mixture models to reconstruct individual state sequences. These models could obviously take into account other types of discrete such as physiological parameters, but they might be extended to continuous state variables (for instance, temperature). Besides, recent developments in the field of conditional random fields might also be investigated to take into account more complex time dynamics or continuous state sequence.

## 15 **Unsupervised extraction and analysis of movement patterns**

A major contribution of the proposed Bayesian framework lies in the non-subjective and unsupervised thus exhaustive categorization of individual movement patterns. 37 different patterns of habitat use, with 20 patterns accounting for more than 80 % of the sample, were identified within the processed sample set. The treatment of the same data set was performed according to a supervised classification in a previous study (Daverat and Tomas *in press*). This categorization exploited only six classes and failed in describing with precision the repertoire of behaviour of the eels. Those six classes had been defined *a priori* from the visual inspection of all the plots of the individual Sr:Ca series and from results for other population found in the literature which was not very robust. Previous work indeed mainly relies on such supervised classification with a view to testing *a priori* hypotheses on patterns of habitat use.

In most cases, some individual patterns did not fit to these *a priori* hypotheses (Kotake et al. 2005; Tsukamoto and Arai 2001; Tzeng et al. 2002) and were withdrawn from the analysis. On the contrary, the robust and unsupervised categorization of movement patterns greatly improves the investigation of unknown populations thanks to the direct determination of the diversity of the patterns of habitat use and of the associated proportions.

The diversity of habitat use chronologies reported here is consistent with habitat use patterns found for *A. anguilla* in the Baltic sea (Limburg et al. 2003; Tzeng et al. 1997) as well as those of other temperate eel species such as *Anguilla japonica* (Tsukamoto and Arai 2001; Tzeng et al. 2002), *A. rostrata* (Cairns et al. 2004; Jessop et al. 2002; Morrison et al. 2003), as well as *A. australis* and *A. dieffenbachii* (Arai et al. 2004). The present work confirmed the existence of eels resident of their capture site (about 28%). Besides, as suggested by eels named “transients” (Tsukamoto & Arai 2001) or “nomads” in (Daverat et al. 2004; Daverat et al. 2005), a significant number of eels (about 72%) changed habitats once or more. Chronologies with multiple transitions between the river and the estuary found for some eels collected in the estuary might be interpreted as an absence of movement under the seasonal fluctuations of river flows into the estuary. However, the estimated habitat-related models account for these seasonal fluctuations and Sr:Ca distributions are well discriminated whatever the season. This makes unlikely the reconstruction of such mislabelled sequences, so that the hypothesis of patterns with multiple or seasonal movements was confirmed in the present study.

As a by-product of the proposed approach, statistical descriptors of the fish movements between habitats, such as the distribution of the transition time from one habitat to another or the distributions of the time spent in a given habitat after a transition, were computed over the whole sample set. This resulted in a huge gain in analysis time and power, compared to previous methods that required to retrieve the information individually from each Sr:Ca

series. The analysis of transitions revealed that movements between two different habitats were not as frequent as the residence in the same habitat along the fish life. The same result was obtained for studies using mark recapture techniques (Jellyman et al. 1996; Morrison and Secor 2003) and telemetry (Parker 1995) that found that most yellow eels tend to adopt a resident behaviour. Transitions are rare temporal events along the fish life which make them difficult to observe directly. In this study, transitions between two different habitats decreased as the age of the eels increased. Similar results were obtained for *Anguilla japonica* (Tzeng et al. 2002) and *A. rostrata* (Morrison et al. 2003) as well as *A. anguilla* (Daverat and Tomas *in press*). Analysis of transitions also revealed that most eels spend less than one year in one habitat before changing again.

The unsupervised categorization framework enlarges the scope of possible analysis in terms of fish ecology. Further developments of the analysis of eel habitat use could consist in comparing individual parameters (size at age, age at maturity), of the different habitat use patterns. Such analysis could also focus on a specific stage of the fish (age or size class). More generally it provides a powerful tool to assess the relative efficiency of individual tactics in terms of fitness. At a broader scale, the proposed approach demonstrates the actual interest in exploiting advanced processing techniques to fully exploit the rich potential of individual biological archive, such as fish otoliths, to characterize individual life traits. A wide range of applications for the analysis of individual life traits might be stated as such a Bayesian reconstruction of the time series of a state sequence from a set of chemical and structural measures.

## Annex I: EM parameter estimation for Gaussian mixture models

The estimation of the parameters of the Gaussian mixture models is carried out according to the maximum likelihood (ML) criterion. It resorts to the following maximization issue:

$$5 \quad \left\{ \hat{\pi}_H, \hat{\Theta}_H, \hat{\sigma}_H \right\}_{H \in \{R, E, M\}} = \arg \max_{\{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}}} \prod_i p(y_i | a_i, s_i, \{\pi_H, \Theta_H, \sigma_H\}_{H \in \{R, E, M\}})$$

To solve for this maximization issue, we use the EM (Expectation-Maximization) algorithm (Bishop 1995). Let us denote by  $x_i$  the variable stating that the  $i^{\text{th}}$  sample is issued from habitat  $x_i$ . The EM algorithm iterates until convergence two steps. At iteration  $k$ , the E-step computes the posterior likelihood  $p(x_i | y_i, a_i, s_i, \{\pi_H^k, \Theta_H^k, \sigma_H^k\}_{H \in \{R, E, M\}})$  of assigning  $x_i$  to habitat  $H$  given

10 the observed measure  $y_i$  and the current estimate of the mixture parameters  $\{\pi_H^k, \Theta_H^k, \sigma_H^k\}_{H \in \{R, E, M\}}$ :

$$p(x_i = H_1 | y_i, a_i, s_i, \{\pi_H^k, \Theta_H^k, \sigma_H^k\}_{H \in \{R, E, M\}}) = \frac{\pi_{H_1}^k p(y_i | a_i, s_i, \Theta_{H_1}^k, \sigma_{H_1}^k)}{\sum_{H_2 \in \{R, E, M\}} \pi_{H_2}^k p(y_i | a_i, s_i, \Theta_{H_2}^k, \sigma_{H_2}^k)}$$

Given the posterior probabilities, the M-step aims at updating model parameters  $\{\pi_H^{k+1}, \Theta_H^{k+1}, \sigma_H^{k+1}\}_{H \in \{R, E, M\}}$ . To simplify the notations, let us denote by  $\tau_{iH}$  the posterior

15  $p(x_i | y_i, a_i, s_i, \{\pi_H^k, \Theta_H^k, \sigma_H^k\}_{H \in \{R, E, M\}})$ . The new priors  $\{\pi_H^{k+1}\}_{H \in \{R, E, M\}}$  are updated as:

$$\pi_H^{k+1} = \frac{1}{N} \sum_i \tau_{iH},$$

The new model parameters  $\Theta_H^{k+1} = (m_H, \lambda_H, \mu_H)$  are estimated as the solution of the following weighted least-square problem, where the weights are give by posterior likelihoods  $\tau_{iH}$ :

$$20 \quad \Theta_H^{k+1} = \arg \min_{\Theta} \left[ \sum_i \tau_{iH} (y_i - Z_i^t \Theta)^2 \right],$$

where  $Z_i$  is the vector defined by  $[1 \ a_i \ s_i]$ . This weighted linear regression leads to:

$$\Theta_H^{k+1} = \left[ \sum_i \tau_{iH} Z_i^t Z_i \right]^{-1} \sum_i \tau_{iH} y_i Z_i ,$$

Then, the updated standard deviations  $\{\sigma_H^{k+1}\}_{H \in \{R, E, M\}}$  are computed from the weighted average of the squared residual error  $r_{iH} = y_i - Z_i^t \Theta_H^{k+1}$  with respect to the prediction issued from the current model estimate:

5

$$\pi_H^{k+1} = \sqrt{\frac{1}{N} \sum_i \tau_{iH} r_{iH}^2} .$$

## References

- Arai, T., Kotake, A., Lokman, P.M., Miller, M.J. and Tsukamoto, K. 2004. Evidence of different habitat use by New Zealand freshwater eels *Anguilla australis* and *A. dieffenbachii*,  
5 as revealed by otolith microchemistry. Mar. Ecol. Prog. Ser. **266**: 213-225.
- Awad, A.M. 1996. Properties of the Akaike information criterion. Microelectron. Reliab. **36**:  
457-464.
- Berg, R. 1985. Age determination of eels, *Anguilla anguilla* (L.): comparison of field data with  
otolith ring patterns. J. Fish Biol. **26** . 537-554.
- 10 Bishop, C. 1995. Neural Networks for Pattern Recognition. Oxford University Press, Oxford.
- Cairns, D.K., Shiao, J.C., Iizuka, Y., Tzeng, W.N. and MacPherson, C.D. 2004. Movement  
patterns of American eels in an impounded watercourse, as indicated by otolith  
microchemistry. N. Am. J. Fish. Man. **24**: 452-458.
- Campana, S.E. 1999. Chemistry and composition of fish otoliths: Pathways, mechanisms and  
15 applications. Mar. Ecol. Prog. Ser. **188**: 263-297.
- Campana, S.E. and Casselman, J.M. 1993. Stock Discrimination Using Otolith Shape-Analysis.  
Can. J. Fish. Aquat. Sci. **50**: 1062-1083.
- Campana, S.E., Chouinard, G.A., Hanson, J.M., Frechet, A. and Bratley, J. 2000. Otolith  
elemental fingerprints as biological tracers of fish stocks. Fish. Res. **46**: 343-357.
- 20 Cardinale, M., Doering-Arjes, P., Kastowsky, M. and Mosegaard, H. 2004. Effects of sex,  
stock, and environment on the shape of known-age Atlantic cod (*Gadus morhua*) otoliths.  
Can. J. Fish. Aquat. Sci. **61**: 158-167.
- Daverat, F. and Tomas, J. *in press*. Tactics and demographic attributes of the European eel  
(*Anguilla anguilla*): the case study of the Gironde watershed (Southwest France). Mar. Ecol.  
25 Prog. Ser.
- Daverat, F., Elie, P. and Lahaye, M. 2004. Microchemistry contribution to a first approach to  
the diversity of life histories of eels from the lower part of the Gironde Garonne Dordogne  
watershed. Cybium. **28**: 83-90.
- Daverat, F., Tomas, J., Lahaye, M., Palmer, M. and Elie, P. 2005. Tracking continental habitat  
30 shifts of eels using otolith Sr/Ca ratios: validation and application to the coastal, estuarine and  
riverine eels of the Gironde-Garonne-Dordogne watershed. Marine and Freshwater Research.  
**56**: 619-627.
- Elsdon, T.S. and Gillanders, B.M. 2005. Consistency of patterns between laboratory  
experiments and field collected fish in otolith chemistry : an example and applications for  
35 salinity reconstructions. Mar. Freshwater Res. **56**: 609-617.
- Fablet, R. 2005. Semi-local extraction of ring structures in images of biological hard tissues:  
application to the Bayesian interpretation of fish otoliths for age and growth estimation. Can.  
J. Fish. Aquat. Sci.
- Jellyman, D.J., Glova, G.J. and Todd, P.R. 1996. Movements of shortfinned eels, *Anguilla*  
40 *australis*, in Lake Ellesmere, New Zealand: Results from mark-recapture studies and sonic  
tracking. N. Z. J. Mar. Freshwater Res. **30**: 371-381.
- Jessop, B.M., Shiao, J.-C., Iizuka, Y. and Tzeng, W.-N. 2002. Migratory behaviour and habitat  
use by American eels *Anguilla rostrata* as revealed by otolith microchemistry. Mar. Ecol.  
Prog. Ser. **233**: 217-229.
- 45 Kawakami, Y., Mochioka, N., Morishita, K., Tajima, T., Nakagawa, H., Toh, H. and Nakazono,  
A. 1998. Factors influencing otolith strontium/calcium ratios in *Anguilla japonica* elvers.  
Environ. Biol. Fish. **52**: 299-303.
- Kotake, A., Okamura, A., Yamada, Y., Utoh, T., Arai, T., Miller, M.J., Oka, H. and Tsukamoto,  
K. 2005. Seasonal variation in the migratory history of the Japanese eel *Anguilla japonica* in  
50 Mikawa Bay, Japan. Mar. Ecol. Prog. Ser. **293**: 213-225.

- Kraus, R.T. and Secor, D.H. 2003. Response of otolith Sr:Ca to a manipulated environment in young American eels. *Am. Fish. Soc. Symp.* **33**: 79-85.
- Kraus, R.T. and Secor, D.H. 2004. Incorporation of strontium into otoliths of an estuarine fish. *J. Exp. Mar. Biol. Ecol.* **302**: 85-106.
- 5 Limburg, K.E., Svedang, H., Elfman, M. and Kristiansson, P. 2003. Do stocked freshwater eels migrate ? Evidence from the Baltic suggests "yes". *In* Biology, management and protection of catadromous eels. American Fisheries Society, Symposium 33, Bethesda. pp. 275-284.
- Martin, G.B. and Thorrold, S.R. 2005. Temperature and salinity effects on magnesium, manganese, and barium incorporation in otoliths of larval and early juvenile spot *Leiostomus*
- 10 *xanthurus*. *Mar. Ecol. Prog. Ser.* **293**: 223-232.
- McCullagh, P. and Nelder, J.A. 1989. Generalized linear models. Chapman and Hall, Londres.
- Milton, D.A. and Chenery, S.R. 2003. Movement patterns of the tropical shad hilsa (*Tenualosa ilisha*) inferred from transects of <sup>87</sup>Sr/<sup>86</sup>Sr isotope ratios in their otoliths. *Can. J. Fish. Aquat. Sci.* **60**: 1376-1385.
- 15 Morrison, W.E. and Secor, D.H. 2003. Demographic attributes of yellow-phase American eels (*Anguilla rostrata*) in the Hudson River estuary. *Can. J. Fish. Aquat. Sci.* **60**: 1487-1501.
- Morrison, W.E., Secor, D.H. and Piccoli, P.M. 2003. Estuarine habitat use by Hudson river American eels as determined by otolith Strontium:Calcium ratios. *In* Biology, management and protection of catadromous eels. American Fisheries Society, Symposium 33, Bethesda.
- 20 pp. 87-100.
- Mounaix, B. and Fontenelle, G. 1994. Anguilles estuariennes et fluviales: apports de l'otolithométrie. *Bull. Fr. Peche Piscic.* **335**: 67-80.
- Murphy, K. 2004. Conditional Random Field (CRF) Toolbox for Matlab. [www.cs.ubc.ca/~murphyk/Software/CRF/crf.html](http://www.cs.ubc.ca/~murphyk/Software/CRF/crf.html).
- 25 Nabney, I. 2001. Netlab: Algorithms for Pattern Recognition. Springer.
- Nelson, C.S., Northcote, T.G. and Hendy, C.H. 1989. Potential use of oxygen and carbon isotopic composition of otoliths to identify migratory and non-migratory stocks of the New Zealand common smelt : a pilot study. *N. Z. J. Mar. Freshwater Res.* **23**: 337-344.
- Parker, S.J. 1995. Homing ability and home range of yellow phase American eels in a tidally
- 30 dominated estuary. *J. Mar. Biol. Assoc. U.K.* **75**: 127-140.
- Rabiner, L.R. 1989. Tutorial on hidden Markov models and selected applications in speech recognition. *Proceedings of the IEEE.* **77**: 257-286.
- RDevelopmentCoreTeam. 2005. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.
- 35 Silber, J.H., Rosenbaum, P.R. and Ross, R.N. 1995. Comparing the contribution of groups of predictors : which outcomes vary with hospital rather than patient characteristics ? *J. Amer. Statist. Assn.* **90**: 7-18.
- Thorrold, S.R. and Shuttleworth, S. 2000. In situ analysis of trace elements and isotope ratios in fish otoliths using laser ablation sector field inductively coupled plasma mass spectrometry.
- 40 *Can. J. Fish. Aquat. Sci.* **57**: 1232-1242.
- Thorrold, S.R., Jones, C.M. and Campana, S.E. 1997. Response of otolith microchemistry to environmental variations experienced by larval and juvenile Atlantic croaker (*Micropogonias undulatus*). *Limnol. Oceanogr.* **42**: 102-111.
- Troade, H., Benzinou, A., Rodin, V. and Le Bihan, J. 2000. Use of deformable template for
- 45 two-dimensional growth ring detection of otoliths by digital image processing : Application to plaice (*Pleuronectes platessa*) otoliths. *Fish. Res.* **46**: 155-163.
- Tsukamoto, K. and Arai, T. 2001. Facultative catadromy of the eel *Anguilla japonica* between freshwater and seawater habitats. *Mar. Ecol. Prog. Ser.* **220**: 265-276.
- Tsukamoto, K., Nakai, I. and Tesch, W.V. 1998. Do all freshwater eels migrate? *Nature.* **396**:
- 50 635.



- Tzeng, W.N. 1996. Effects of salinity and ontogenetic movements on strontium:calcium ratios in the otoliths of the Japanese eel, *Anguilla japonica*. J. Exp. Mar. Biol. Ecol. **199**: 111-122.
- Tzeng, W.N. and Tsai, Y.C. 1994. Changes in otolith microchemistry of the Japanese eel, *Anguilla japonica*, during its migration from the ocean to the rivers of Taiwan. J. Fish Biol. **45**: 671-683.
- 5 Tzeng, W.N., Wu, H.F. and Wickstrom, H. 1994. Scanning electron microscopic analysis of annulus microstructure in otolith of European eel, *Anguilla anguilla*. J. Fish Biol. **45**: 479-492.
- 10 Tzeng, W.N., Severin, K.P. and Wickstrom, H. 1997. Use of otolith microchemistry to investigate the environmental history of European eel *Anguilla anguilla*. Mar. Ecol. Prog. Ser. **149**: 73-81.
- Tzeng, W.N., Shiao, J.C. and Iizuka, Y. 2002. Use of otolith Sr:Ca ratios to study the riverine migratory behaviors of Japanese eel *Anguilla japonica*. Mar. Ecol. Prog. Ser. **245**: 213-221.

Factor	LR Chisq	Df	Pr(>Chisq)
Habitat	43744	2	< 2.2e-16
Season	31	1	2.397e-08
Age	1984	15	<2.2e-16

**Table 1: Anova table for the linear model.**

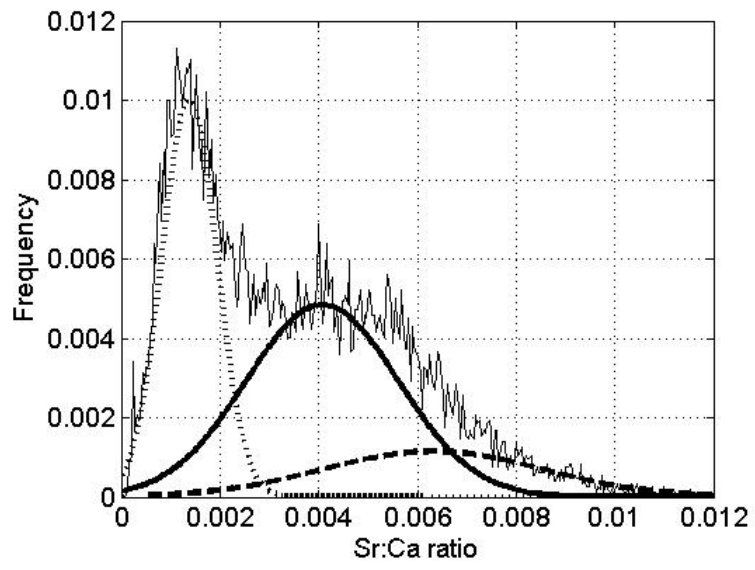


Fig. 1.

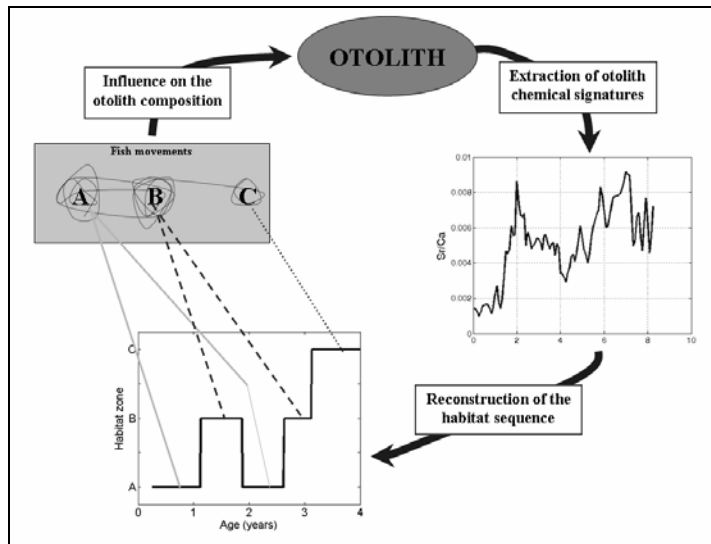


Fig. 2.

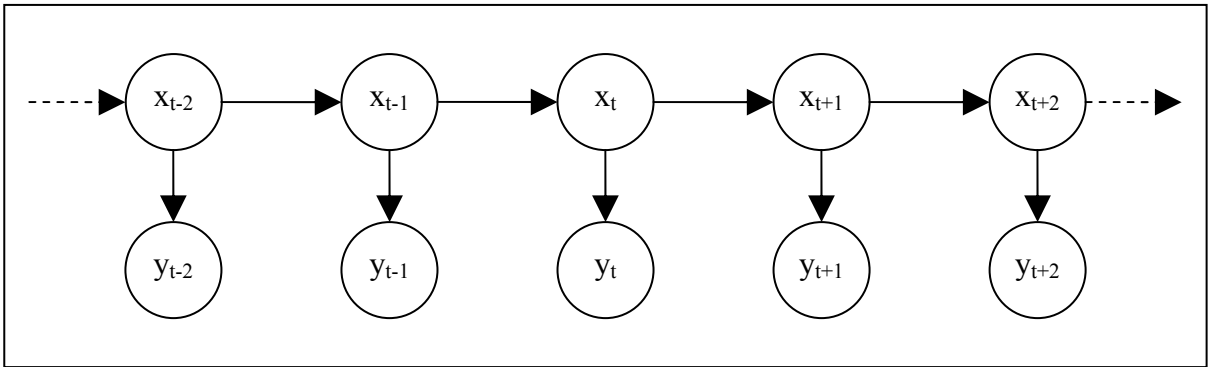


Fig. 3.

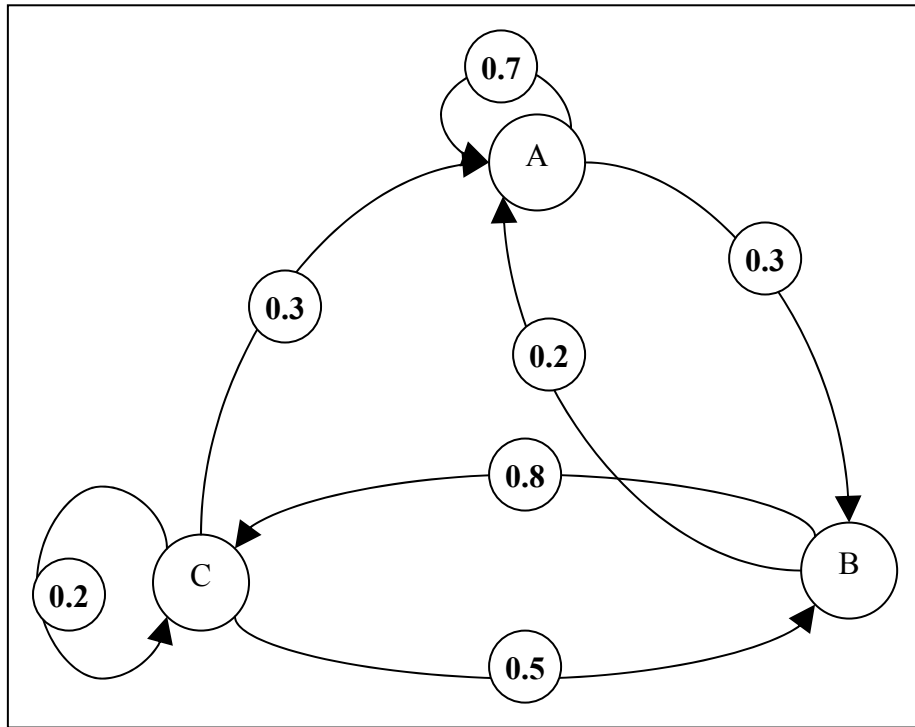


Fig. 4

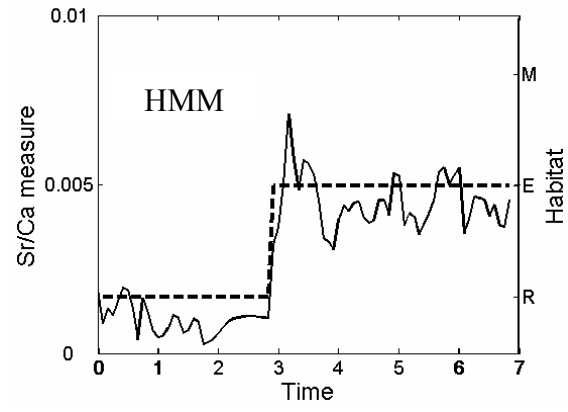
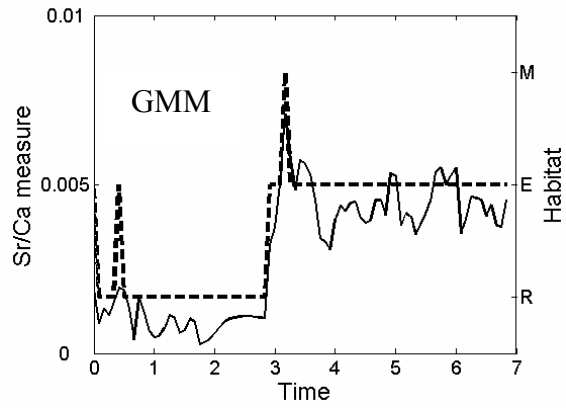
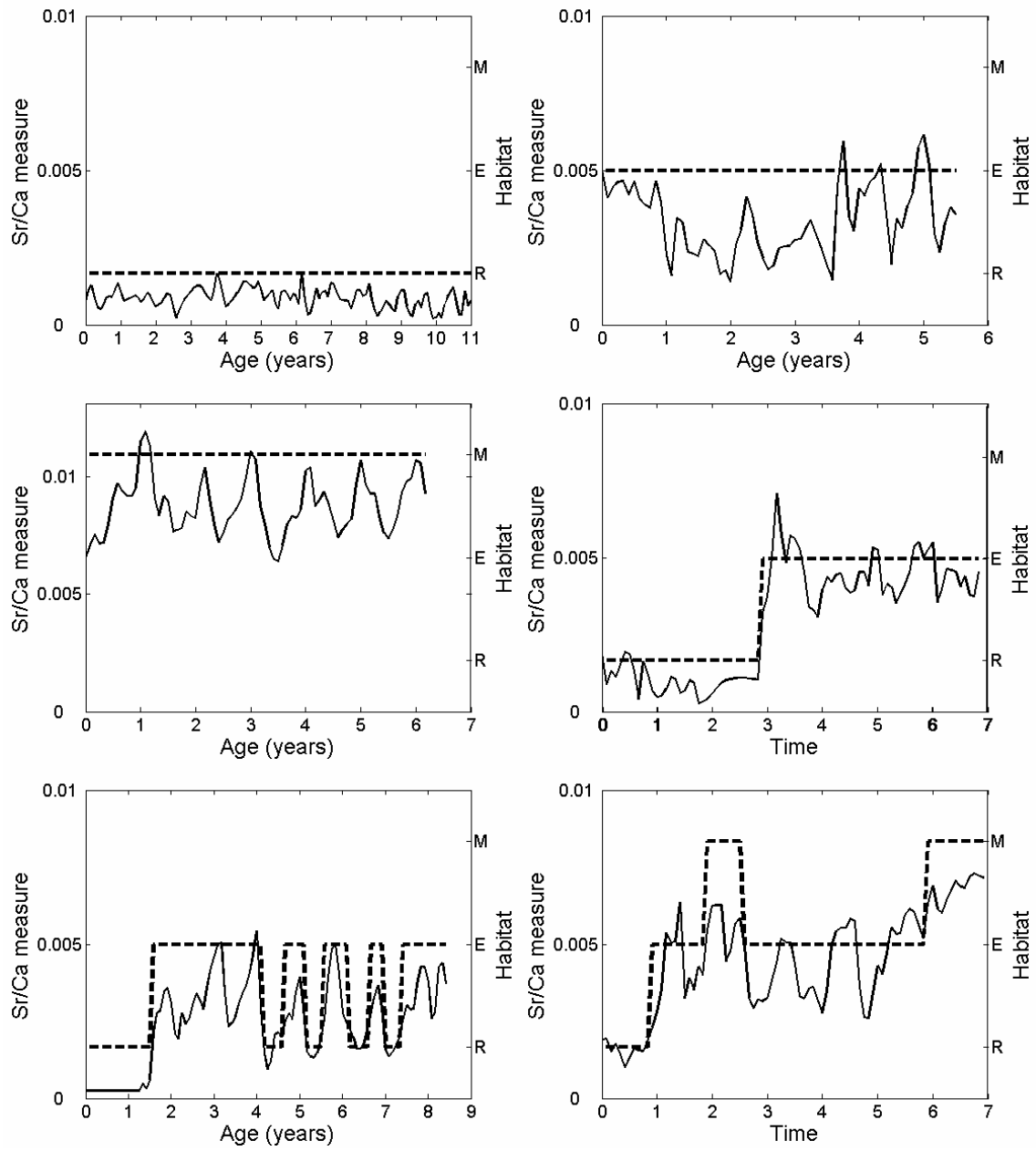


Fig. 5



5

Fig. 6



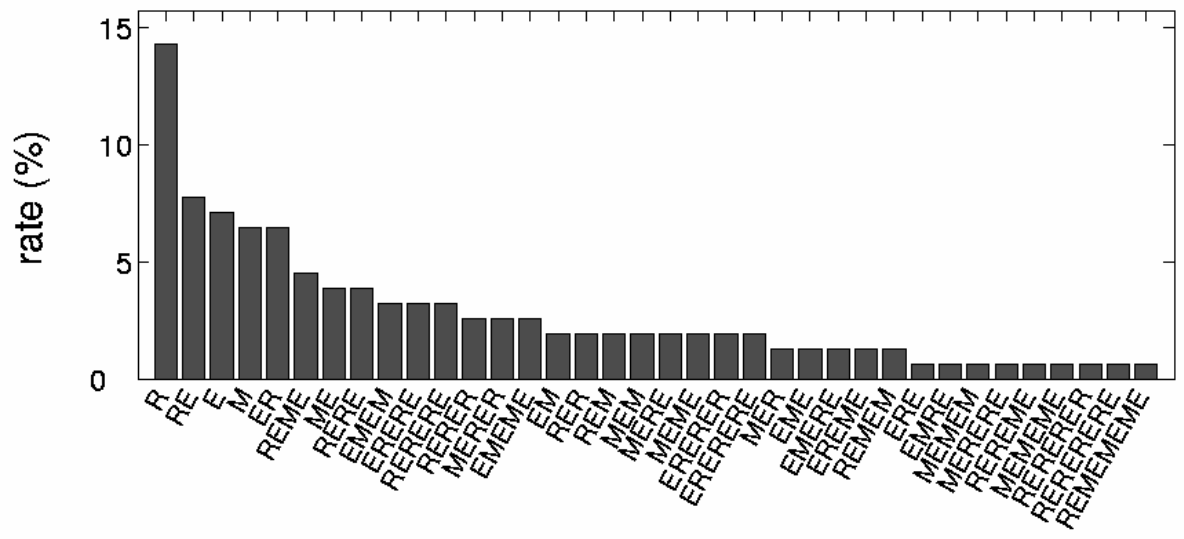


Fig. 7

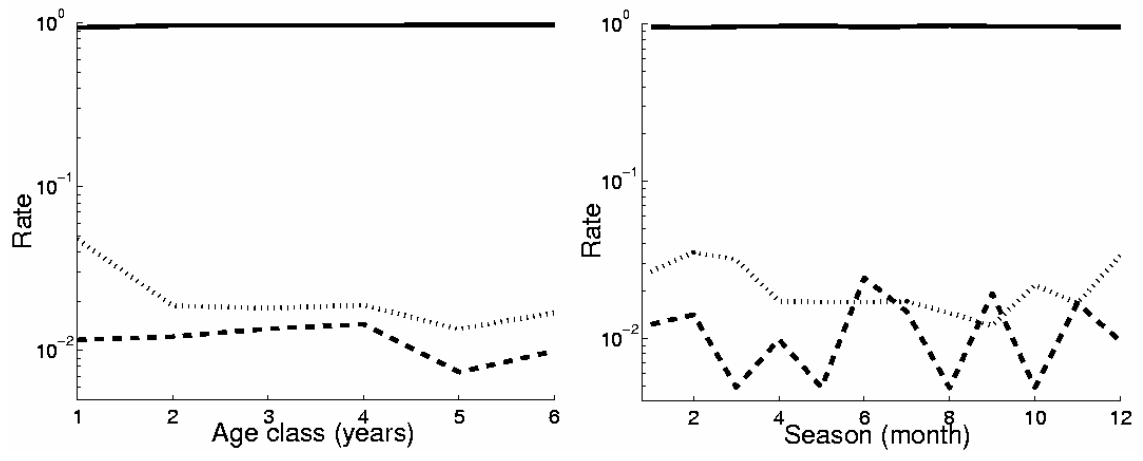


Fig. 8

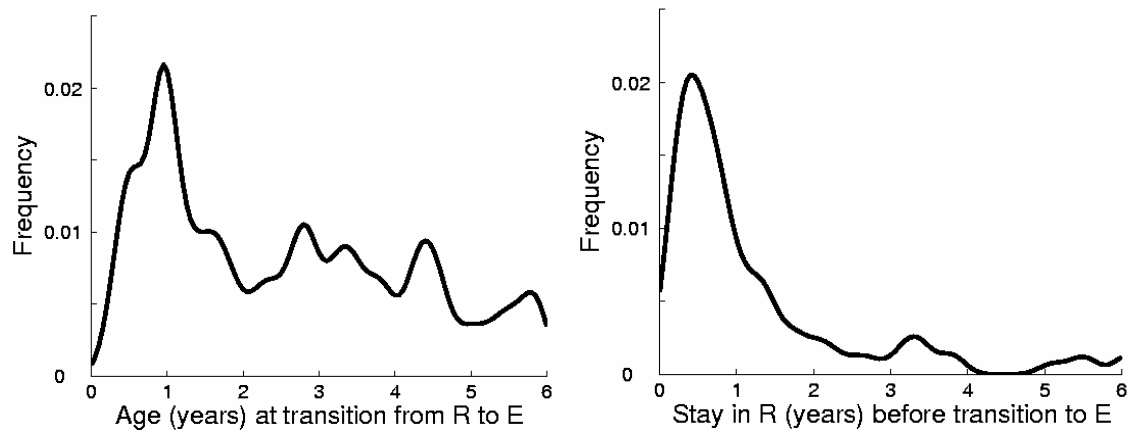


Fig. 9

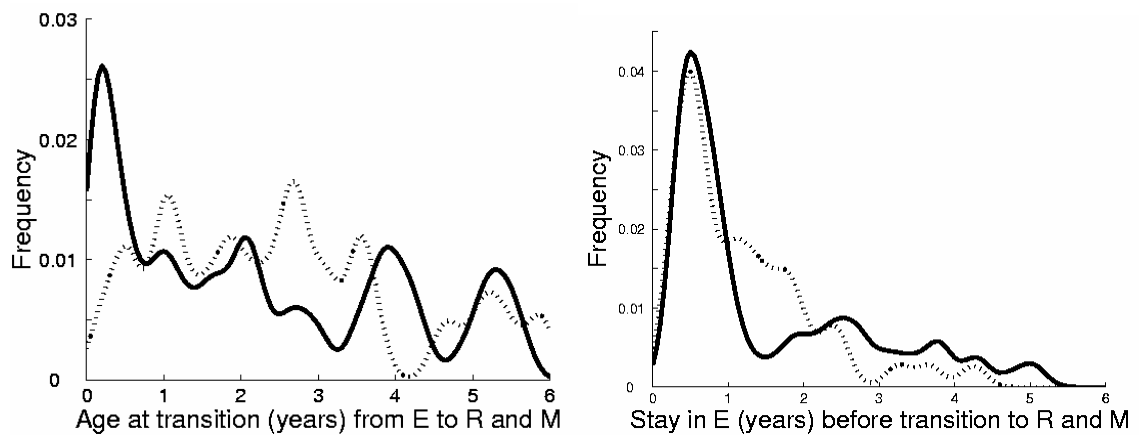


Fig. 10.

Figure 1: Distribution of Sr:Ca values and associated fitted Gaussian modes.

5 Figure 2: Principle of the reconstruction of the time habitat sequence from the acquired Sr:Ca measures spatially sampled along a growth axis of the otolith.

10 Figure 3: Graphical representation of the Gaussian hidden Markov models: the arrows indicate the conditional dependencies  $x_t|x_{t-1}$  for the temporal dynamics and the conditional dependencies  $y_t|x_t$  for modeling the likelihood of measure  $y_t$  given state  $x_t$ .

15 Figure 4: Illustration of the characteristics of the transition matrix which specifies the temporal dynamics of the state variable for a model involving four states A, B, C and D: graphical representation of this transition matrix (the arrows illustrate the conditional dependencies between these states with associated likelihood with the associated transition probabilities). As illustrated, the probability to leave each state is 1. Some transitions may be forbidden, i.e. associated with a null probability: for instance, transitions from B to A. A particular case of transitions is the one corresponding to staying in the current state (for instance, transitions A to A or C to C).

20 Figure 5: Habitat use pattern issued from a given Sr:Ca series as obtained with the Gaussian mixture model (GMM) (left panel) or the hidden Markov model (HMM) (right panel).

25 Figure 6: Examples of estimated patterns of habitat for four individual eels consistent with a residency in the estuary (upper left panel), or with a residency in a freshwater habitat (upper right panel), or a shift of habitat from the river to the estuary (bottom left panel) or multiple movements between the river and the estuary (bottom right panel) X axis, estimated age (years), plain line, Sr:Ca values, dash line, Hidden Markov Model estimation.

30 Figure 7: Analysis of the movement patterns: frequencies of the movement patterns extracted from the estimated habitat sequences.

35 Figure 8: Proportions (Y axis, log scale) of instantaneous transitions from the estuarial states to one of the three states (namely, river (dots), estuary (plain), marine (dash)) as a function of age or month.

Figure 9: Distribution of ages at transition from the river to the estuary and distribution of the river habitat use duration anterior to the transition.

40 Figure 10: Distribution of ages at transition from the estuary to the river (plain line) or to the marine habitat (dots) (left panel) and distribution of the estuary habitat use duration anterior to the transition (right panel).