Ecological Modelling
July 2010, Volume 221, Issue 15, Pages 1757-1769
http://dx.doi.org/10.1016/j.ecolmodel.2010.04.005

Identifying fishing trip behaviour and estimating fishing effort from VMS data using Bayesian Hidden Markov Models

Youen Vermard[a, b, *], Etienne Rivot[b], Stéphanie Mahévas[c], Paul Marchal[a] and Didier Gascuel[b]

[a] IFREMER, Channel and North Sea Fisheries Department, 150 Quai Gambetta, BP 699, 62321 Boulogne s/mer, France
[b] AGROCAMPUS OUEST, UMR 985 Ecologie et Santé des Ecosystèmes, Ecologie Halieutique, 65 rue de St-Brieuc, CS 84 215, 35 042 Rennes Cedex, France
[c] IFREMER, Fisheries and Ecological Modeling Department, Rue de l'Ile d'Yeu, BP 21105, 44311 Nantes Cedex 03, France

*: Corresponding author : Y. Vermard, Tel.: +33 3 21 99 56 86; fax: +33 3 21 99 56 01, email address : Youen.Vermard@ifremer.fr

Abstract:

Recent advances in technologies have lead to a vast influx of data on movements, based on discrete recorded position of animals or fishing boats, opening new horizons for future analyses. However, most of the potential interest of tracking data depends on the ability to develop suitable modelling strategies to analyze trajectories from discrete recorded positions. A serious modelling challenge is to infer the evolution of the true position and the associated spatio-temporal distribution of behavioural states using discrete, error-prone and incomplete observations. In this paper, a Bayesian Hierarchical Model (HBM) using Hidden Markov Process (HMP) is proposed as a template for analyzing fishing boats trajectories based on data available from satellite-based vessel monitoring systems (VMS). The analysis seeks to enhance the definition of the fishing pressure exerted on fish stocks, by discriminating between the different behavioural states of a fishing trip, and also by quantifying the relative importance of each of these states during a fishing trip. The HBM approach is tested to analyse the behaviour of pelagic trawlers in the Bay of Biscay. A hidden Markov chain with a regular discrete time step is used to model transitions between successive behavioural states (e.g., fishing, steaming, stopping (at Port or at sea)) of each vessel. The parameters of the movement process (speed and turning angles) are defined conditionally upon the behavioural states. Bayesian methods are used to integrate the available data (typically VMS position recorded at discrete time) and to draw inferences on any unknown parameters of the model. The model is first tested on simulated data with different parameters structures. Results provide insights on the potential of HBM with HMP to analyze VMS data. They show that if VMS positions are recorded synchronously with the instants at which the process switch from one behavioural state to another, the estimation method provides unbiased and precise inferences on behavioural states and on associated movement parameters. However, if the observations are not gathered with a sufficiently high frequency, the performance of the estimation method could be drastically impacted when the discrete observations are not synchronous with the switching instants. The model is then applied to real pathways to estimate variables of interest such as the number of operations per trip, time and distance spent fishing or travelling.

Keywords: Bayesian Hierarchical Models; Hidden Markov Model; State-space model; VMS; Fleet behaviour; Fishing effort

INTRODUCTION

Recent advances in technologies have lead to a vast influx of data on movements of animals or fishing boats, opening new horizons for future analyses of movements, trajectories and behaviours to address fundamental (e.g. analyzing foraging behaviours) or applied (e.g. analyzing fishing strategy) issues. However, most of the potential interest of tracking data depends on the ability to develop suitable modelling strategies to analyze trajectories from discrete recorded positions. Thus, a serious modelling challenge this paper seeks to address is to infer the evolution of the true position and the associated spatio-temporal distribution of behavioural states using discrete, error-prone and incomplete observations. The interest of inferring on animal spatial distribution and behaviour has been recently addressed in several studies (Barraquand and Benhamou 2008, Jonsen et al. 2005, Patterson et al. 2008). Accounting for spatial and seasonal characteristics of fishing activities is essential for reliable stock assessments and realistic forecasting models for management purposes (Booth, 2000; Babcock et al., 2005; Pelletier and Mahévas, 2005). A fine scale spatio-temporal description of fishing behaviours, effort and catches provides insights for a better understanding of both the spatio-temporal dynamics of fish resources (Bertrand et al., 2004; Poos and Rijnsdorp, 2007), and the impact of fishing pressure on marine ecosystems (Smith and Wilen, 2003; Rijnsdorp et al., 1998; Mills et al., 2007). The exploration of alternative management measures is another field of application. For instance, understanding fishermen response to management measures is critical to anticipate the effect of management strategies (Vermard et al., 2008) and simulation tools for management scenario testing require a spatial description of vessels' dynamics (Mahévas and Pelletier, 2004).

Classical methods to analyse fishing effort are based on data derived from fishermen declarations (log-books). In the North-East Atlantic, fishing effort data are often recorded as days at sea and reported at the scale of the ICES[1] rectangle (30' in latitude and 1° in longitude). However, both the metric used and the reporting scale are too coarse for accurately estimating fishing effort, and may give a misleading picture of its actual structure (Rijnsdorp et al., 1998). Deriving a fine scale spatio-temporal distribution of fishing activity and fishing effort requires analysing the activity of fishing vessels at sea, which can typically be divided between travelling/steaming time, searching time, fishing time and handling time. Log-books are not designed to provide information that can be used for such a fine scale description of fishing trips. Distinguishing between these different phases or behaviours would have two main benefits. First, it would enable us to improve the definition of the effective fishing effort, i.e. the pressure that is actually exerted by fishing units on harvested stocks. Second, although the different phases of the fishing trip may overlap (skipper searching for fish schools when the crew is processing the fish already caught), all these activities usually result in distinct revenues and costs. From an economic point of view, it is then important to be able to quantify the duration of these different phases (Pelletier et al., 2009).

Recent advances in technologies have lead to a vast influx of data on movements of fishing boats, thereby opening new horizons for future analysis. In 1998, the European Commission (EU) introduced legislation to monitor European fishing vessels for security control and enforcement purposes using a satellite-based Vessel Monitoring System (VMS). From 1st January 2005, all vessels over 15 m in length are required to transmit their position at interval of 2 hours or less. These data provide a discrete, more or less regular record of the vessels position. It is therefore thought that VMS data are a potentially valuable source of information to understand spatial and temporal dynamics of fishing activity, fishing effort allocation, costs and revenues, and of biological impacts of fisheries.

[1] International Coucil for the Exploration of the Sea

However, VMS data basically consist in sequentially recorded positions, and do not directly indicate whether a vessel is fishing or not. Most of the potential use of VMS data then depends upon our ability in interpreting these records to accurately distinguish travelling/steaming from searching and fishing behaviour during boat trips. Building statistical tools to analyse VMS data hence remains challenging.

Classical analyses of VMS data use vessel's speed and sometimes vessel's direction rules (speed between two positions and turning angle between two segments) to identify trawling and steaming behaviour. These analyses assume that boats steaming will mostly follow a straight line at a high speed and boats fishing are characterized by a more erratic trajectory and a low speed. Several authors (e.g. Kourti et al., 2005; Murawski et al., 2005; Harrington et al., 2007; Mills et al., 2007) have proposed methods that necessitate strong hypotheses to be set a priori. In particular, the angle and speed characterizing the different behavioural states have to be specified a priori. Moreover, such methods are appropriate when the travelling and fishing speeds are very different and when the boats are not practicing different fishing activities with different fishing speeds. Instead of assuming a linear interpolation of the track, Hintzen et al. (2009) used the cubic Hermite method to improve its description. They however classified the position recording based on a speed level set *a priori*. Bertrand et al., (2005; 2007) proposed to describe the movement on its own through random walk based on Lewy trajectories. The method was applied to characterize and quantify the entire movement of foragers, and it is not designed to separate out fishing and travelling time.

These methods are not fully satisfying and inferring the evolution of the true but hidden position and behavioural state of fishing vessels from available (discrete, error-prone and incomplete) recorded VMS positions reveals an exiting challenge.

In this paper, we propose the Bayesian hierarchical modelling framework as a general template to analyse fishing vessel trajectories from VMS data. Bayesian hierarchical models (BHM) (Clark 2005; Cressie et al. 2009) using hidden markov models (HMM) have been proposed recently has a valuable framework to deliver the analytical basis for a synthesis on individual movements (Patterson et al., 2008). The framework was successfully applied to analyse movement data of animals from imprecise or incomplete survey data. Morales et al. (2004) applied BHM to elk movements and found associations between different behaviours (encamped or exploratory mode) and habitat type, respectively woodland and agricultural habitat. Jonsen et al. (2005) analyse the foraging behaviours of seals through Bayesian SSM of remotely sensed movements. Jonsen (2006) and Jonsen et al. (2007) applied the approach to analyse the behaviour and trajectory of leatherback turtles.

But to our best knowledge, BHM has never been applied to model fisheries behaviour. By contrast with models usually developed to study movement from VMS data and to distinguish fishing from steaming, it is not necessary to specify a priori the value of the speed and turning angles characterizing each behaviour. In theory, this approach can also accommodate missing position records which are quite usual in VMS data. However, if the BHM approach theoretically offers some flexibility to deal with complex spatio-temporal models (Cressie et al. 2009), its practical implementation for analysing VMS data remains challenging, and the aim of this paper is to provide a first investigation of the potential of HBM with hidden Markov processes to analyse VMS data.

The approach is developed in three steps. First, the main intuition of modelling fishing boats behaviours through hidden Markov process in continuous or discrete time is pointed out. Second, a specific model with three behavioural states (fishing, steaming and stopping) within a discrete time Markovian framework is developed. The performance of the Bayesian estimation method is assessed through a simulation-estimation approach. Several contrasted scenarios were played to assess how different data configurations impact the estimations. In a third step, the framework was applied to the French pelagic fishery of the Bay of Biscay. This fleet is targeting various pelagic species (e.g., Anchovy, Sardine, Tuna, Horse Mackerel) (Vermard et al., 2008) and can operate at a large scale going from the whole Bay of Biscay to the Channel. It has been affected by a severe crisis from 2005 following the anchovy closure. Given the possible stock recovery and re-opening of the fishery, some management measures such as spatial closures or effort reduction are envisaged. We discuss the extent to which improving fishing effort metrics via our approach could contribute to develop the scientific rationale supporting these management measures.

MATERIAL AND METHODS

VMS data

Vessel Monitoring System (VMS) was introduced as part of the European Common Fishery Policy. It is applied to boats over 24 meters since 01/01/2000 (CE No 686/97), to boats over 18 meters since 01/01/2004 and to boats over 15 meters since 01/01/2005 (CE No 2244/2003). Vessels are monitored by system using Inmarsat, Euteltracs or Argos systems. Position (accuracy around 500 m ; FA0, 1998), time (accuracy = 1 sec.; FAO 1998) and, since 2005, heading and instantaneous speed are recorded for each vessel. These data are recorded at a time step inferior to two hours. However, time intervals between two emissions are often not regular, or the boat position can even be unknown for hours because of lack of satellite coverage, breakdown or stops in the emission system. The irregularity and the gaps in the available time series can blur the information contained in these VMS data and complicates the identification of states, speeds and boats pathways.

A hidden Markov process for modelling fishing boat behaviours

This section explains the key intuition of the modelling framework and seeks to point out the main methodological issues addressed in the paper.

Bayesian Hierarchical Models with Hidden Markov Process

The approach consists in coupling an hypothetical and hidden (non observed) mechanistic model of individual movements including stochasticity, to an observation model including observation error, which gives the probability of obtaining a particular observation conditional on the true position and behavioural state. The hidden process of individual movement is modelled through Markovian transitions between different behavioural states, related to the movement process. The succession of the behavioural states forms the so called hidden (not observed) Markov chain. Typically, distributions for speed and turning angles are associated with each behavioural state. At each time step, the approach enables one to estimate the true position, the probability to be in a particular state (behavioural mode), and the process model parameters (e.g. mean speed and turning angles). The Bayesian framework has several advantages for deriving inferences in such complex models. First, the Bayesian setting offers the opportunity to integrate multiple sources of information through data and informative priors. Second, inferences come in the form of posterior probability distributions, which fully describe uncertainty. Third, Monte Carlo simulation methods and associated softwares provide efficient techniques to estimate the posterior distribution even for such kind of models with complex hierarchical structure (Lunn et al. 2009).

Markov process in continuous time as a general template

The main intuition of the model consists in considering the successive alternation of the fishing boats behaviours as a hidden Markov process (MP). MPs in continuous time provide a general template for modelling movement behaviour and in particular fishing boat behaviour. Let us suppose a MP in continuous time, denoted S_t, taking its value in a discrete states space of size k, with possible states in $\{1,\ldots,k\}$. In our application, S_t will denote the state of fishing boats at time t, and S_t will take values in $\{1,2,3\}$, the three possible states being steaming, fishing or stopping. In a first order homogeneous continuous MP (also called memory-less; the future state of the system is influenced only by its current state and not by the past), the amount of time T_i the process stays in state i before shifting to another state is random with an exponential distribution with rate λ_i (λ_i>0) depending upon the current state i (Karlin and Taylor 1975; Ross 1996). The greater the rate λ_i, the smaller the mean time spent in state i before switching. Once a shift happens, one needs to define the direction (the state) in which the shift will occur. The probability to shift from the current state i to an other state j ($j \neq i$) is denoted p_{ij} ($\sum_{j=1}^{k} p_{i,j}=1$ for all $i,$ and $p_{ii}=0$ because we are working conditionally upon a shift happens). Hence, the probabilities p_{ij}'s and the rates λ_i's capture the stochastic structure of a continuous Markovian process.

MP in discrete time can be considered as a simplification of Markov process in continuous time in the sense that the amount of time T_i the process stays in state i before shifting to another state are random but take discrete values (an entire number of time steps). Instead of an exponential distribution, the distribution of the T_i*'s* are geometric. Such models can alternatively be viewed as Markov process in discrete and regular time step Δt (we can define $\Delta t = 1$ without any loss of generality). The Markov chain is now viewed at any discrete time step $t = 1,2,\ldots,n$, and not at the switching instants as before. The process is entirely defined by the $k \times k$ stochastic matrix $P = (p_{i,j})$ where $p_{i,j}$ is the probability to shift from state i to state j

between two discrete times t and $t+1$ (with $\sum_{j=1}^{k} p_{i,j} = 1$ for all i, and $p_{i,i}$ can be non null as the system might well stay in the same state i between two instants t and $t+1$).

For instance, the Figure 1 sketches the behaviour of a fishing boat switching between two states steaming (state 1) and fishing (state 2). This behaviour can be modelled in a continuous (A) or in a discrete with regular time steps (B) framework. Through a MP in discrete time with regular time steps, the switching events arise at the end of a given time step, the amount of time spent on each behavioural state is a multiple of the time step duration. Through a MP in continuous time, the amounts of time spent in both states are random, and the mean amount of time spent at fishing is smaller than the amount of time spent at steaming, what corresponds to $\lambda_1<\lambda_2$. The impact of approximating a continuous MP by a discrete MP is not an issue addressed in this paper. Rather, the article is focussed on the performance of the estimation method when the system is observed at discrete time.

Fig. 1 near here

Drawing inference from observations acquired at discrete time

The MP for the states of the system mimics the dynamic of the successive behaviours of a fishing boat, which is not directly observed. The observations one are willing to use are the successive positions registered from VMS data, which are acquired at a rather regular time steps because of the VMS device.

Let us suppose a first (ideal) situation in which observations about the state of the system are acquired precisely at the instants at which the system switches from one state to another. Irrespective of the framework used for the hidden MP for fishing boat behaviour (e.g. either continuous or discrete time with regular time step), such a situation can be qualified as *data-*

rich in the sense that the available observations are informative about the hidden MP. The observations are the VMS positions at each switching instant. The time interval between two observations provides direct information about the amount of time spent in the current state. Two successive observations provide information about the speed of the boat, and hence about the behavioural state of the boat between the time interval considered, and three successive positions provide information about the change of direction and are in turn also informative about the behavioural state.

However, such a situation is not realistic, as the instants at which VMS data are acquired do not have any chance to match with instants at which boats switch from one behavioural state to another. Indeed, GPS devices are routinely programmed to send an emission at roughly regular time step (say of 1 hour), totally independently from the rhythm of the fishing activity. Hence, irrespective of the framework (continuous or discrete time with regular time steps), deriving inferences about the behavioural states of the boats from observations acquired at a discrete (roughly) regular time step independently from the rhythm of the fishing activity becomes challenging. For instance, the Figure 1 illustrates a case where the observations are acquired at regular time steps, no matter the switching points between two different behaviours. If the data are acquired with a rather low time frequency (say 1 hour for instance), then short fishing operation (say about 20' such as the one corresponding to T_6) will be hardly identified. By contrast, if the frequency of the data acquisition increases (see the effect of additional information in Fig. 1), the performance of the estimation method should increase. For instance, the identification of the operation T_6 (Fig. 1) should be improved by increasing the acquisition rate.

Here, by using a simplified discrete time Markov process framework for the dynamic of fishing boat behaviour, we propose to address the following questions through a simulation method: 1) What is the performance of the estimation method in the ideal situation where observations are available at each switching instants between two behaviours? 2) What is the performance of the method when the observations are available at instants which do not correspond to the switching instants between different behaviours? 3) What is the performance when the frequency of the data acquisition increases/decreases?

These questions are addressed through a specific model with three behavioural states for the fishing boats developed in the following section.

Specific state-space model with a hidden Markov chain with three behavioural states

Process model

Fig. 2 near here

The model is organized following a hierarchical structure (Fig. 2). At the top of the structure, constant parameters control the hidden Markov chain that mimics the sequence of behavioural states and the associated movement throughout time. At the bottom of the structure, the observations are defined conditionally upon the true positions. The movement model was built on discrete time step (in accordance with the data, this time step represents 1 hour). Inspiring from Jonsen et al. (2005), the process model was built to deal with three different states of the boats ("Stopping", "Steaming" and "Fishing"). Using the terminology defined in Morales et al. (2004), the model was defined as a "Triple-switch" model. The movement parameters are indexed by each behavioural mode.

Markovian model for behaviour transitions

At each time step the behavioural mode of the boat is denoted S_t (Fig. 2). A first order homogeneous Markovian model mimics the probabilistic switch between the three behavioural states from one step to another, given the current behavioural state. The transition kernel is defined by a 3x3 matrix of switching probability considered as constant over time, denoted P, with the $p_{i,j}$'s the probability of moving from behavioural state i to behavioural state j (1 is behavioural state "Fishing", 2 is "Steaming" and 3 "Stopping").

Movement model

The movement is also defined on a discrete time step. The movement equation defines the location of the boats over regular time intervals given the previous state and location and the current behavioural mode. Let us denote X_t (a two-dimensional vectors of longitude and latitude) the position of the boat at each time step t. Conditionally upon the behavioural node S_t, the next location X_{t+1} is built using the displacement D_{t+1} computed from the speed and turning angle associated with the current behavioural state S_t assuming a straight line travel between X_t and X_{t+1}. The process error term ε_{t+1} being bivariate Normal with a variance-covariance matrix σ_p^2:

(1) $$X_{t+1} = X_t + D_{t+1} + \varepsilon_{t+1} \quad \text{with} \quad \varepsilon_{t+1} \sim N(0, \sigma_p^2)$$

with the displacement D_{t+1} vector defined as:

(2) $$D_{t+1} = V_t \cdot T_t \cdot U_t$$

$U_t = \dfrac{D_t}{\|D_t\|}$ is an orthonormal vector that gives the direction of the previous movement. Both V_t and T_t depend upon the behavioural state of the boat during the current time step $t \rightarrow t+1$. V_t (a scalar) is the speed of trawler movement given the trawler is in state S_t during the movement D_{t+1}. Speeds are embedded within a hierarchical structure such that at each time step t, V_t is drawn in a prior with unknown mean that depend upon the current behavioural state S_t:

(3) $V_t \mid S_t \sim N(\mu_{S=S_t}, {\sigma^2}_{S=S_t})$

When the boat is at behavioural mode "Stopping", speed is set at 0, no displacement is made. T_t is the transition matrix at time t with mean turning angle θ_t that defines the rotational component of the movement, such that $T_t \cdot U_t$ is the new direction after turning angles :

(4) $$T_t = \begin{bmatrix} \cos(\theta_t) & -\sin(\theta_t) \\ \sin(\theta_t) & \cos(\theta_t) \end{bmatrix}$$

Following Morales et al. (2004) and Eckert et al. (2008), turning angles are distributed a priori as a Wrapped-Cauchy distribution (Fisher 1993). W-Cauchy distributions are embedded within a hierarchical structure such that at each time step t, θ_t is drawn in W-Cauchy distribution with concentration parameter ρ that depends upon the current behavioural state S_t. Following Eckert et al. (2008), location parameters of W-Cauchy were set to 0 (μ_θ=0):

(5) $\theta_t \mid S_t \sim \text{Wrapped-Cauchy}(\rho_{S_t}, \mu_\theta = 0)$

When the boat is at behavioural mode "Stopping" a directional vector U_t is built randomly to be able to compute the next displacement.

Observation model

The observation equation links the unobservable states of the boats predicted by the process model above to the available data (i.e. the recorded position). In the most favourable case where a recorded position y_t (two-dimensional vector) is available at each time step t, the observation equation is modelled using a bivariate normal distribution with variance-covariance matrix σ_0^2 fixed a priori (variance=0.1 and covariance=0) to mimic the low error structure of the location observation (FA0, 1998) :

(6) $$y_{t+1} = X_{t+1} + \omega_{t+1} \quad \text{with} \quad \omega_{t+1} \sim Normal(0, \sigma_0^2)$$

The observation equation (6) was adapted to cope with observations that are not synchronous with the time step of the state process. Following Jonsen et al. (2005), let us denote $t+\delta_t$ the time at which an observation is available between t and $t+1$, δ_t corresponding to a fraction of an entire time step. Assuming a straight line travel between X_t and $X_{t+\delta t}$, the unobserved position of the boat at time $t+\delta_t$, $Z_{t+\delta t}$ and the associated observation errors are defined as follows:

(7a) $$Z_{t+\delta_t} = X_t + \delta_t \cdot (X_{t+1} - X_t) = X_t + \delta_t \cdot D_{t+1}$$

(7b) $$y_{t+\delta_t} = Z_{t+\delta_t} + \omega_{t+\delta_t} \quad \text{with} \quad \omega_{t+\delta_t} \sim Normal(0, \sigma_0^2)$$

This observation equation (7) allows for handling several values of δ_t in a given time step.

Bayesian estimation

Prior

For all unknown parameters, we used rather vague priors based on some reasonable constraints (Table 1).

The mean speed while steaming was drawn in a uniform distribution (with large bounds), and the mean speed while fishing was considered a priori smaller than during steaming. The mean concentration parameter for the W-Cauchy distribution of turning angles while fishing was drawn in a uniform distribution (with appropriate bounds) and the mean concentration parameter while steaming was considered a priori higher than while fishing to mimic the a priori hypothesis that the movement while fishing is more erratic than while steaming. Standard deviation for speed were drawn in uniform distributions with large bounds. The probabilities in the transition matrix P were drawn a priori in rather vague Dirichlet distributions (Congdon, 2001), that is a multivariate generalization of the beta distribution and widely used to model proportions. $p_{2,3}$ and $p_{3,2}$ were assigned very low values to mimic the prior idea that the corresponding transitions are practically impossible. The matrix of variance-covariance σ_p^2 was drawn in a rather vague Whishart distribution (Congdon, 2001).

Table 1 near here

Indetermination due to interpolation and missing data

Equations (7a,b) are needed to cope with time-lags between the switching instants of the Markov process and the instants at which VMS positions are available. The interpolation defining the state $Z_{t+\delta t}$ is simple in theory. However, it is not so easy to cope with in practice as it may lead to a lack of statistical identifiability. In practice, it may lead to a model indetermination. The Figure 3 illustrates that different true paths (defined by the true positions $\{X_t\}$) may correspond to the same interpolated positions $\{Z_t\}$ and therefore to the same sequential observations $\{y_t\}$. In the inferential reasoning, such kind of configuration for the observed recorded positions $\{y_t\}$ may in turn lead to a statistical indetermination of the true path $\{X_t\}$ and therefore to the associated movement parameters.

The problem has its maximum intensity when the time-lag is 0.5, and becomes worth when missing data occur. To minimize interpolation problems during estimation when missing data occur, lag-time surrounding missing values were artificially set to zero and the end of all simulated paths were fixed by adding five successive emissions at the same location simulating a “Stop” at the end of each path.

Fig. 3 near here

Simulation-Estimation approach

Objectives

To assess the sensitivity of the model to the data structure (lack of contrast in speed and turning angles between the various behavioural modes, time-lags between the switching instants of the Markov process and the instants at which VMS positions are recorded) a simulation-estimation (SE) approach was first carried out. The chart flow of the SE approach has 4 steps: *i*) Simulate pathways with known parameters; *ii*) Given a true pathway, simulate different scenarios for observed locations with progressive degradation of the information; *iii*) Use the HBM framework to estimate true pathways, behavioural states and underlying parameters; *iv*) Measure the performance of the estimation method by comparing the Bayesian estimation of the unknowns with the values used for the simulations.

Scenarios

12 contrasted scenarios were tested (Table 2) to investigate how the quality of the inferences varies with several data configurations. Computation being very time-consuming it was not possible to undertake a factorial experiment considering all possible combinations of configurations for the Markov process model and the observation model. Consequently a few scenarios were carefully selected that illustrated effects of particular parameters, so as to be the most informative on the likely performance of the method and sequentially addressing different questions following the two main axes: *i*) Movement process: Is it possible to accurately identify behavioural states ("Steaming" and "Fishing"), even when the contrast between the associated movements becomes weaker?; *ii*) Observation process: in real data set, recorded VMS positions are necessarily recorded with time lags between the instants at which the boats switch from one behavioural state to another and the recording instants. Moreover, missing data exist (long periods without any recorded position). Several scenarios were

played to assess whether such kind of data configurations enable to derive accurate inferences, and to assess the impact of increasing the frequency of the observations.

For all scenarios, a pathway of 100 time steps (approx 4 days) was simulated as follow. First, a sequence of behavioural states was simulated following the Markovian model with transition matrix P. The switching probabilities were set as:

$$P=\begin{bmatrix} 0.7 & 0.2 & 0.1 \\ 0.6 & 0.4 & 0 \\ 0.5 & 0 & 0.5 \end{bmatrix}$$

Then, at each time step, conditionally upon the behavioural states at time t, a speed V_t and a turning angle θ_t were drawn in their distribution associated with the behavioural states, and the displacement was computed deterministically from eq. (1)-(2). A sequence of observations was then computed following the observation equation (7a,b). The Figure 4 presents the simulated pathway for scenarios 1 and 2.

Fig. 4 near here

Scenario 1. The first scenario was built to be as close as possible from the speeds and turning angles distributions observed in real data from the French pelagic fishery in the bay of Biscay. First, Average "Fishing" speed was set to 4 knots (σ_{Fi}=1.5) and average "Steaming" speed to 10 knots (σ_{St}=1.5). Angles were drawn in a Wrapped-Cauchy distribution with concentration parameter equal to 0.2 and 0.5 for "Fishing" and "Steaming" respectively. Observation are recorded at each time sep of the MP without time lag.

Scenario 2. This scenario mimics a case with more distinct movements characteristics behaviours between Fishing and Steaming, the distributions of angles and speed (mean for "Fishing" = 4 knots and "Steaming" = 10 knots) being more constrained around the means

(σ=0.1 for speed for both “Fishing” and “Steaming”, concentration parameters for the Wrapped-Cauchy distribution for turning angles equal to 0.5 and 0.9 for “Fishing” and “Steaming” respectively). Observations were recorded at each time step of the process and without time lag.

Scenarios 3-8. These scenarios are based on the reference scenario 1, but the observations are blurred by adding time-lag between switching instants of the process and observations (Table 2). Equations (7a,b) are used, with specified values for the time-lags δ_t. For instance, in the scenario 8, a constant value $\delta_t = 0.1$ is used at all time steps. Several levels of lag-time and structure of the lag-time were tested: scenarios 3-5 are characterized by different values of constant time-lags, whereas scenarios 6-8 tested different configurations of random time-lags.

Scenarios 9-11. These scenarios are based on the scenario 8, but missing data were introduced in the recorded positions to reproduce sequences of missing values typically observed in real datasets. Several levels of missing values were introduced, going from 5% of the time steps of the pathway to 20%.

Scenarios 12. This scenario aims at assessing the impact of raising the level of information in a scenario where the model is not able to provide reliable estimates of the trajectory. It is based on the scenario 7, but observations were simulated at a higher frequency (3 observations per time step).

Table 2 near here

Bayesian estimation from simulated data and performance of the estimation method

The following methods were used to evaluate he performance of the estimation method. Concerning the speed, we compute the relative bias which is $(E(\theta|y) - \theta_{true})/\theta_{true}$, where $E(\theta|y)$

is the expected mean of the posterior distribution and θ_{true} the mean of the distribution of speed used for simulation (see Table 2). We also computed $V(\theta|y)^{1/2}$ to measure the Bayesian uncertainty around the estimates. Concerning the inferences on the behavioural states, we assessed the percentage of behavioural states which are correctly predicted along each 100 steps pathway. At each time step *t*, the posterior credibility of each of the three behavioural states is readily obtained from posterior inferences. The behavioural state *S* is attributed a posteriori to the time step *t* if *S* is the most credible a posteriori of the three possible states, and the state is said well predicted if the state attributed a posteriori matches with the simulated state.

Application to observed VMS data

The model was then applied to real pathways of pelagic trawlers from which VMS data could be made available. A pathway of 398 time steps, containing only 9 missing data (1 missing data is considered to occur when the interval between two successive emissions is approx. 2 hours) and for which VMS emission are obtained at very regular time intervals (~1 hour) was used as an example of application. At that period of the year, the fishery is essentially targeting sea bass with trawling sequences usually longer than 1 hour (around 5-6 hours and up to 8 hours (Morizur et al. 1996)). This allows us to suppose that the emission with frequency of about 1 hour are rather informative with regards to the succession of behavioural states. Posterior inferences on behavioural states were used to extract relevant measures of the fishing effort. Posterior probability distributions of, e.g., the distance covered during steaming or fishing, or the number of fishing operations per trip, were also computed. To allocate a behavioural mode to each position, the same procedure than in the SE approach was used, but a threshold probability Π_{min} was introduced: the behavioural state *S* is attributed to the time step *t* if *S* is the most credible of the three possible states and if the posterior probability of *S* is greater of equal to the threshold Π_{min}. No behavioural state ("unknown state") is allocated to time step t if none of the three states has a posterior probability greater than Π_{min}. The sensitivity of the classification to the value of the threshold Π_{min} was assessed with values of Π_{min} varying between 0.5 to 1.

Technical details

The estimation was performed using the OpenBUGS software (http://www.mrc-bsu.cam.ac.uk/bugs/) and the BRugs package of R (www.r-project.org) (Lunn et al. 2009). The OpenBUGS software offers huge modelling flexibility. It uses Monte Carlo Markov Chains simulations to provide estimates of the posterior distributions. Three independent MCMC chains with different initialisation points were used. For each chain, the first 20 000 iterations were discarded as an initial burn-in period. Inferences were then derived from the next 30 000 iterations, but only one out of 10 iterations was kept to reduce the MCMC sampling autocorrelation, leading to 3 000 iterations by chain. Hence inferences were derived from a sample of 9 000 iterations proceeded from three chains of 3 000 iterations each. The convergence of all MCMC chains was checked via the Gelman-Rubin diagnostics.

RESULTS

Simulation-Estimation approach

Impact of the similarity of the behavioural state parameters

Comparing the inferences between scenarios 1 and 2 (Fig. 5) highlights that the Bayesian hierarchical model provides very high quality inferences, even in the case where the contrast between the behavioural states (in term of speed and turning angles) is low.

For the reference scenario 1, the model is also able to reproduce the simulated pathway, estimate speed with low bias and uncertainty (Fig. 5) and is able to accurately capture mostly all the behavioural states (Table 3). In the scenario 2, the two states "Fishing" and "Steaming" are characterized by more distinct speed and turning angles distributions than in the scenario 1. Logically, the estimations of all pathways characteristics have very low bias and very low uncertainty (Fig. 5 and Table 3). However, the gain in the quality of the inferences comparing to scenario 1 is only weak.

Fig. 5 near here

Table 3 near here

Introducing time-lags

Comparing scenarios 1 and 3-8 highlights that the inferences are highly sensitive to the introduction of time-lags between the discrete process movements and the recorded observations, and that inferences may rapidly become unreliable if most of the time lags are near 0.5.

Scenarios with time-lags either small or high (scenarios 3, 5, 6 and 8), provide very good estimation of speed (small bias and uncertainty in the estimated speeds) (Fig. 5). By contrast, scenarios where lots of emissions are made in the middle of the time step (scenarios 4 and 7), provide very poor fits with high uncertainty in speed estimates and lots of behavioural modes are not correctly identified (Table 3). Poor capacity to predict behavioural state is linked with a poor fit of the displacement parameters with high uncertainty (Fig. 5). The problem of statistical indetermination anticipated in the Material and Methods section (Fig. 3) becomes critical in the scenario 7 where many observed positions y_t are recorded with time-lags near

0.5. The estimated path X_t and the associated movement parameters are highly uncertain (Fig. 5 and Table 3).

Introducing missing values

Fig 5 and Table 3 show that the more missing values in the pathway (scenarios 9, 10 and 11) the more bias and uncertainty in the estimation of speed and of the behavioural states. However, all behavioural states are not affected in the same proportions. For instance, with 20% of missing data, respectively 85 and 92% of the “Steaming” and “Fishing” positions are correctly identified, but only 57% of the “Stopping” positions are correctly estimated (Table 3). “Stopping” positions which are not correctly identified are confounded with either “Steaming” and “Fishing” positions.

Raising the number of observations

Comparing scenarios 7 and 12 (Fig. 5 and Table 3) shows that increasing the frequency at which VMS positions are gathered drastically increases the performance of the estimation method, even if these observations are not synchronous with the switching instants.

Application to a real dataset

Given the results of the simulation-estimation approach, the real data set that we analysed corresponds to a rather favourable situations (the percentage of missing values is rather low, approx. 2%, and the frequency at which the VMS positions are acquired is shorter than the mean duration of fishing operations). We therefore consider the posterior inferences as rather reliable.

Each true hidden location is identified with a very low level of uncertainty. The behaviour “Stopping” is allocated to some of the time steps, and many of the associated positions match with known geographical locations of harbours in the Bay of Biscay. The other ones are interpreted as stop at sea.

The posterior distributions of the associated movements characteristics, such as speed in each behavioural states, are readily estimated. Fig 6 shows the posterior distributions of speeds while “Steaming” and “Fishing”. During these fishing trips, the estimated mean speed while “Fishing” and “Steaming” are respectively 4 and 11 knots, which is consistent with knowledge issued from previous studies (Morizur et al., 1996). Other interesting indicators can also be readily estimated, such as the time spent in each state (Fig 7) or the distance travelled in each state (Fig 8). It is worth noting that the uncertainty about these estimates is rather low.

Fig. 6 near here

Fig. 7 near here

Fig. 8 near here

The inferences are only weakly sensitive to the threshold value Π_{min} chosen in the allocation rule for the behavioural states (Fig 9). Indeed, assigning behavioural modes using Π_{min}=0.5 leads to similar results that the method consisting in assigning the behavioural states with the highest posterior probability (as in the simulation/estimation approach). This low sensitivity reflects the fact that the behavioural states are identified with little ambiguity: most of the time, one of the three states has a posterior probability which is far greater than the two others.

Fig. 9 near here

DISCUSSION

This paper shows that Bayesian hierarchical models using hidden Markov process are a promising approach to describe boats movements and identify behavioural states during a trip from discrete recorded VMS positions. The method is adapted when the movement can be a priori divided in various modes (Barraquand and Benhamou, 2008). It appears therefore well suited to disentangle the time spent in different behavioural modes during a fishing trip and to analyse fishing behaviour and fishing effort. Here, we investigated the potential of HBM with hidden Markov process to analyse VMS data, using a Markov processes in discrete time for sake of simplification. In particular, our simulation-estimation approach was designed to address questions regarding the performance of the estimation method according to various parameters (synchronism between records and switches between two different states, frequency of observations and missing values). These questions are all relative to the quantity of information provided by the data relative to the process, and can be considered, at least in a first approach, as relatively independent from the modelling framework (discrete or continuous) chosen for the hidden Markov process. Hence, a hidden Markov process in discrete time, which is easier to program for Bayesian inferences, was used as a first approach.

The simulation-estimation approach provides an analysis of the performance of the method, and contributes to evaluate the degree of confidence in the outputs of the model when interpreting results from real data sets. Given the multiple combinations of levels of parameters for the process and the observation model, a few scenarios were selected to illustrate the effect of particular parameter. Results highlighted that when the VMS positions are precisely recorded at the switching instants, the estimation methods performs well, the model being able to reproduce the true pathway, to capture very well the sequence of behavioural modes, and to provide unbiased estimates of the parameters (speed and angles) characterizing the movements in each behavioural mode. The model performs remarkably well even if the behavioural modes are not associated with clearly distinct movements characteristics. However, besides these very uncourageous results, our analysis also pointed out that the estimation performances are drastically impacted when the positions are not recorded synchronously with the switching instants. In this case, reliable inferences can still be obtained if the frequency with which the data are recorded is greater than the frequency with which the process switches from one behavioural mode to another.

The conclusions of the simulation-estimation approach are very insightful regarding the potential use of VMS data to track fishing boats behaviours at a fine temporal and spatial scale. VMS emissions are now routinely gathered at time interval of approximately 1 hour. It is worth noting that these data should reveal relatively non informative if the fishery under concern has fishing operations with mean duration shorter than 1 hour (*e.g.* trawling duration of 20' for instance). By contrast, if the fishing operations are much longer (*e.g.* about 2 hours such as the purse seine tuna fishery and up to 6 hours for some trawling fisheries such as the pelagic fishery while targeting sea bass as example in this paper), then VMS emission every hour could be successfully used to efficiently track the succession of behaviours. These quite intuitive results put forward the critical question of the frequency at which VMS data should be acquired, in order to give some feed back to managers that fix the acquisition time period for different fishing boats and fishing activities practiced. Our very first conclusion is that a 1-hour frequency is certainly too long to be able to correctly capture behaviours for all fishing boats and all fishing activities.

This first analysis opens several perspectives for future work. As stated in the Material and Methods section, Markov processes in continuous time constitutes the general template for modelling fishing trips behaviours. Indeed, the amounts of time spent in each behaviour certainly take values in the continuous time line. Future research should be undertaken to propose a continuous-time MP framework to analyse VMS data. Using a Markov process in continuous time would improve the switching time identification and in the same time the underlying parameters' estimation (speed and turning angles). It is worth noting that inferring the switching points (and the associated behaviours) of a continuous Markov process from discrete recorded positions is certainly a more difficult problem than working with a discrete Markov process. Indeed, a mismatch in observation and switch point times caused by random variation in observation timing has different implications than a mismatch caused by random variation in transition times, and may certainly lead to a more complicated inferential problem. Although the effects of random variation in observation times or random variation in switch points may not be a very important distinction in situations where the frequency of observations is much higher than the frequency of possible switch points (or if the discrete MP possible switch points are as frequent the probability of remaining in the same state for multiple time units is quite high), this however, is not the general case. This constitutes an important issue to be addressed in the future.

An other promising perspective would be to integrate in the model the information brought about by onboard observers. Indeed, these data provide us with invaluable information on fishers' behaviour at sea as they record the true sequence of the onboard operations such as fishing, stopping for gear maintenance, searching, steaming. Fishing trips for which onboard observers data are available could be used to improve the definition of the different fishing behaviours and their succession in time and space, or these data could be used in a first analysis to derive informative priors distributions for further analysis. More generally, the Bayesian framework is promising as it allows to integrating multiple sources of information, including expertise, in the modelling framework. For instance, the first order homogeneous

Markovian hypothesis is very strong and could be relaxed to integrate the idea that the behaviours at each instant depends upon the whole history of the fishing trip from the departure of the boats. Also, using the spatial coordinate of all the harbours where boats are potentially landing their harvest could certainly help improving the identification of "stopping" behaviour. The framework could be further improved by including covariates such as maps of the sea bed or primary productions.

Despite the limitations and all the perspective to improve the method, this study provides some insights on how VMS data could be used to characterize effort allocation during a fishing trip. Since 1998 and the beginning of VMS recording, a large amount of data concerning boats operating with different kind of gears, targeting different species in distinct areas have been registered. The diversity of the fishing activities operated requires a flexible method to accommodate a wide range of fishing behaviours. To add to the diversity of the underlying processes, trajectories can be observed throughout various emission systems (Inmarsat, Argos). Our model may be applied to evaluate quantitatively the different stages of fishing trips. Of particular importance for fisheries management is the share of a boat trip that is dedicated to fishing. More generally, this share is one of the behavioural component of fishing that determines the effective fishing effort. Of course, other components have to be taken into account to accurately estimate this effective fishing effort such as, for instance, the efficiency of the research time or the exchange of information between fishermen (Millischer et Gascuel 2006). From that point of view, the analysis of VMS data is step forward in the understanding and quantitative characterisation of fishing behaviour.

Enhancing fishing effort metrics is also particularly important when assessing the impact of fishing on the seabed (Mills, 2007), the effort attraction around Marine Protected Areas (Murawski, 2005) and even fish distribution (Bertrand, 2005). Improving the description of fishing effort would also positively impact the reliability of catch rates as stock abundance indicators (Marchal et al., 2006).

VMS data should, in the future, greatly benefit studies on effort allocation and fishers' behaviour. The statistics derived from these approaches could then be used to compute the effective fishing time and the spatial and temporal patterns of fishing activity. These descriptors could then serve as direct inputs for stock assessment (for instance in calibrating VPA on effort data) and for existing bio-economic modelling frameworks (e.g. ISIS-Fish (Mahévas and Pelletier, 2004; Drouineau et al., 2006; Pelletier et al., 2009, In Press), TEMAS (Ulrich et al., 2002 and 2007) or FLR (Kell et al., 2007)) to improve the modelling of fishery systems. These indicators may also be of direct value for management and monitoring purposes. It is for example important to distinguish between fishing and steaming when establishing Marine Protected Areas that can potentially be crossed by boats because of its location, either between fishing areas, or between the home harbour and fishing grounds. Being able to distinguish fishing from travelling is also important, in the context of input-based management, to adjust fishing effort limits to management objectives.

ACKNOWLEDGEMENTS

The work was funded through the CAFE project of the European Union (DG-Fish, contract no. 022644) and the Région Bretagne, for which support we are very grateful. We are also indebted to fishers, who kindly provided their VMS data on a voluntary basis and people from the French Fisheries Information System at IFREMER. The authors thank Marie-Pierre Etienne, AgroParis Tech, ENGREF, Paris, and Emily Walker and Nicolas Bez (IRD Sète) for helpful comments and discussions and the two anonymous referees for their relevant comments that have greatly improved the paper.

REFERENCES

Babcock, E.A., Pikitch, E.K., McAllister, M.K., Apostolaki, P., Santora, C. (2005) A perspective on the use of spatialized indicators for ecosystem-based fishery management through spatial zoning. *ICES J. Mar. Sci.* **62,** 469-476.

Barraquand, F., Benhamou, S. (2008) Animal movement in heterogeneous landscapes: Identifying profitable places and homogeneous movement bouts. *Ecology* **89,** 3336-3348.

Bertrand, S., Bertrand, A., Guevara-Carrasco, R., Gerlotto, F. (2007) Scale-invariant movements of fishermen: the same foraging strategy as natural predators. *Ecological Applications* **17,** 331-337.

Bertrand, S., Burgos, J.M., Gerlotto, F., Atiquipa, J. (2005) Lévy trajectories of Peruvian purse-seiners as an indicator of the spatial distribution of anchovt (*Engraulis rigens*). *ICES Journal of Marine Science* **62,** 477-482.

Bertrand, S., Diaz, E., Niquen, M. (2004) Interactions between fish and fisher's spatial distribution and behaviour: an empirical study of the anchovy (Engraulis ringens) fishery of Peru. *ICES Journal of Marine Science* **61,** 1127-1136.

Booth, A.J. (2000) Incorporating the spatial component of fisheries data into stock assessment models. *ICES J. Mar. Sci.* **57,** 858-865.

Clark, J.S. 2005 Why environmental scientists are becoming Bayesians. *Ecol Lett*, 8:2-14

Congdon, P. (2001) Bayesian Statistical Modelling. *Wiley & Sons Ldt, Wiley Series in Probability and Statistics***,** 529 p.

Cressie, N., Calder, C.A. Clark, J.S. Ver Hoeff J.M. and Wikle C.K. 2009. Accounting for uncertainty in ecological analysis: the strengths and limitations of hierarchical statistical modelling. *Ecological Applications*, 19(3): 553–570.

Drouineau, H., Mahévas, S., Pelletier, D., Benoît, B. (2006) Assessing the impact of different management options using ISIS-Fish: the French Merluccius merluccius – Nephrops norvegicus mixed fishery of the Bay of Biscay. *Aquatic Living Resources* **19,** 15-29.

Eckert, S.A., Moore, J.E., Dunn, D.C., Sagarminaga Van Buiten, R., Eckert, K.L., Halpin, P.N. 2008. Modeling loggerhead turtle movement in the Mediterranean: importance of body size and oceanography. *Ecological Applications* **18**(2), 290-308.

F.A.0 (1998) Fishing Operations, 1.Vessel Monitoring Systems. *FAO TECHNICAL GUIDELINES FOR RESPONSIBLE FISHERIES* **1**.

Fisher, N.I. 1993. Statistical analysis of circular data. Cambridge University Press, Cambridge, UK.

Harrington, J.J., Semmens, J.M., Haddon, M. (2007) Spatial distribution of commercial dredge fishing effort: application to survey design and the spatial management of a patchily distributed benthic bivalve species. *Marine and Freshwater Research* **58,** 756-764.

Hintzen, N.T., Piet, G.J, Brunel, T. (2009) Improved estimation of trawling tracks using cubic Hermite spline interpolation of position registration data. *Fisheries Research* DOI: 10.1016/j.fishres.2009.09.014

Jonsen, I.D., Flemming, J.M., Myers, R.A. (2005) Robust state-space modeling of animal movement data. *Ecology* **86,** 2874-2880.

Jonsen, I.D., Myers, R.A., James, M.C. (2007) Identifying leatherback turtle foraging behaviour from satellite telemetry using a switching state-space model. *Marine Ecology-Progress Series* **337,** 255-264.

Jonsen, I.D. (2006) Robust hierarchical state-space models reveal diel variation in travel rates of migrating leatherback turtles. *Journal of Animal Ecology* **75,** 1046-1057.

Karlin, S. and Taylor, H.M. 1975. A First Course in Stochastic Processes, Second Edition. Academic Press.

Kell, L.T., Mosqueira, I., Grosjean, P., Fromentin, J.M., Garcia, D., Hillary, R., Jardim, E., Mardle, S., Pastoors, M.A., Poos, J.J., Scott, F., Scott, R.D. (2007) FLR: an open-source framework for the evaluation and development of management strategies. *ICES Journal of Marine Science* **64,** 640-646.

Kourti, N., Shepherd, I., Greidanus, H., Greidanus, H., Alvarez, M., Aresu, E., Bauna, T., Chesworth, J., Lemoine, G., Schwartz, G. (2005) Integrating remote sensing in fisheries control. *Fisheries Management and Ecology* **12,** 295-307.

Lunn, D.J., Thomas, A., Best, N., Spiegelhalter, D. (2000) WinBUGS - A Bayesian modelling framework: Concepts, structure, and extensibility. *Statistics and Computing* **10,** 325-337.

Lunn, D.J., Spiegelhalter, D., Thomas, A., Best, N. (2009). The BUGS project: Evolutin, critique and future directions. *Statist. Med.*, DOI : 10.1002/sim.3680.

Mahévas, S., Pelletier, D. (2004) ISIS-Fish, a generic and spatially explicit simulation tool for evaluating the impact of management measures on fisheries dynamics. *Ecological Modelling* **171,** 65-84.

Marchal, P., Andersen, B.S., Bromley, D., Iriondo, A., Mahévas, S., Quirijns, F., Rackham, B., Santurtun, M., Tien, N., Ulrich, C. (2006) Improving the definition of fishing effort for important European fleets by accounting for the skipper effect. *Canadian Journal of Fisheries and Aquatic Sciences* **63,** 510–533.

Millischer L., Gascuel D. (2006) Information transfers, behaviors of vessels and fishing efficiency: an individual based simulation approach. *Aquatic Living Ressources*, **19**: 1-13.

Mills, C.M., Townsend, S.E., Jennings, S., Eastwood, P.D., Houghton, C.A. (2007) Estimating high resolution trawl fishing effort from satellite-based vessel monitoring system data. *ICES J. Mar. Sci.* **64,** 248-255.

Morales, J.M., Haydon, D.T., Frair, J., Holsinger, K.E., Fryxell, J.M. (2004) Extracting more out of relocation data: building movement models as mixtures of random walks. *Ecology* **85,** 2436-2445.

Morizur, Y., Tregenza, N., Heessen, H.J.L., Berrow, S.D., S., P. (1996) By-catch and discarding in pelagic trawl fisheries. *Contract Report Contract EC DG XIV-C-1, IFREMER, Brest*.

Murawski, S.A., Wigley, S.E., Fogarty, M.J., Rago, P.J., Mountain, D.G. (2005) Effort distribution and catch patterns adjacent to temperate MPAs. *ICES Journal of Marine Science* **62,** 1150-1167.

Patterson, A.T., Thomas, L., Wilcox, C., Ovaskainen, O., Matthiopoulos, J. (2008) Stat-space models of individual animal movement. *Trends in Ecology & Evolution* **23(**2), 87-94.

Pelletier, D., Mahévas, S., Drouineau, H., Vermard, Y., Thébaud, O., Guyader, O., Poussin, B. (2009) Evaluation of the bioeconomic sustainability of multi-species multi-fleet fisheries under a wide range of policy options using ISIS-Fish. *Ecological Modelling* **220**, 1013-1033.

Poos, J.-J., Rijnsdorp, A.D. (2007) The dynamics of small-scale patchiness of plaice and sole as reflected in the catch rates of the Dutch beam trawl fleet and its implications for the fleet dynamics. *Journal of Sea Research* **58,** 100-112.

Rijnsdorp, A.D., Buys, A.M., Storbeck, F., Visser, E.G. (1998) Micro-scale distribution of beam trawl effort in the southern North Sea between 1993 and 1996 in relation to the trawling frequency of the sea bed and the impact on benthic organisms. *ICES J. Mar. Sci.* **55,** 403-419.

Ross, S.M. 1996. Stochastic Processes, 2^{nd} Edition, Wiley, New-Yok.

Smith, M.D., Wilen, J.E. (2003) Economic impacts of marine reserves: the importance of spatial behavior. *Journal of Environmental Economics and Management* **46,** 183-206.
Ulrich, C., Andersen, B.S., Sparre, P.J., Nielsen, J.R. (2007) TEMAS: fleet-based bio-economic simulation software to evaluate management strategies accounting for fleet behaviour. *ICES Journal of Marine Science* **64,** 647-651.
Ulrich, C., Le Gallic, B., Dunn, M.R., Gascuel, D. (2002) A multi-species multi-fleet bioeconomic simulation model for the English Channel artisanal fisheries. *Fisheries Research* **58,** 379-401.
Vermard, Y., Marchal, P., Mahevas, S., Thebaud, O. (2008) A dynamic model of the Bay of Biscay pelagic fleet simulating fishing trip choice: the response to the closure of the European anchovy (Engraulis encrasicolus) fishery in 2005. *Canadian Journal of Fisheries and Aquatic Sciences* **65,** 2444-2453.

Table 1. Prior used for Bayesian estimation (Fi=Fishing, St=Steaming).

Parameters	Prior used
Speed	
Mean speed while Steaming	$\mu_{S=1} \sim Unif(5,20)$
while Fishing	$\begin{cases} \mu_{S=2} = \alpha_v * \mu_{S=1} \\ \alpha_v \sim Beta(2,2) \end{cases}$
Standard deviation for speed	$\sigma_{S=1} \sim Unif(0,10)$
	$\sigma_{S=2} \sim Unif(0,10)$
Turning angles	
Concentration parameter of W-Cauchy	
while Steaming	$\rho_{S=1} \sim Unif(0,1)$
while Fishing	$\begin{cases} \rho_{S=2} = \alpha_p * \rho_{S=1} \\ \alpha_p \sim Beta(1,1) \end{cases}$
Transition matrix P	$(p_{1,1}, p_{1,2}, p_{1,3}) \sim Dirichlet(33,33,34)$
	$(p_{2,1}, p_{2,2}, p_{2,3}) \sim Dirichlet(50,40,0.1)$
	$(p_{3,1}, p_{3,2}, p_{3,3}) \sim Dirichlet(50,0.1,40)$
Variance-covariance for the movement process	$\sigma_p^2 \sim Whishart(\Omega,2)$
	$\Omega = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$

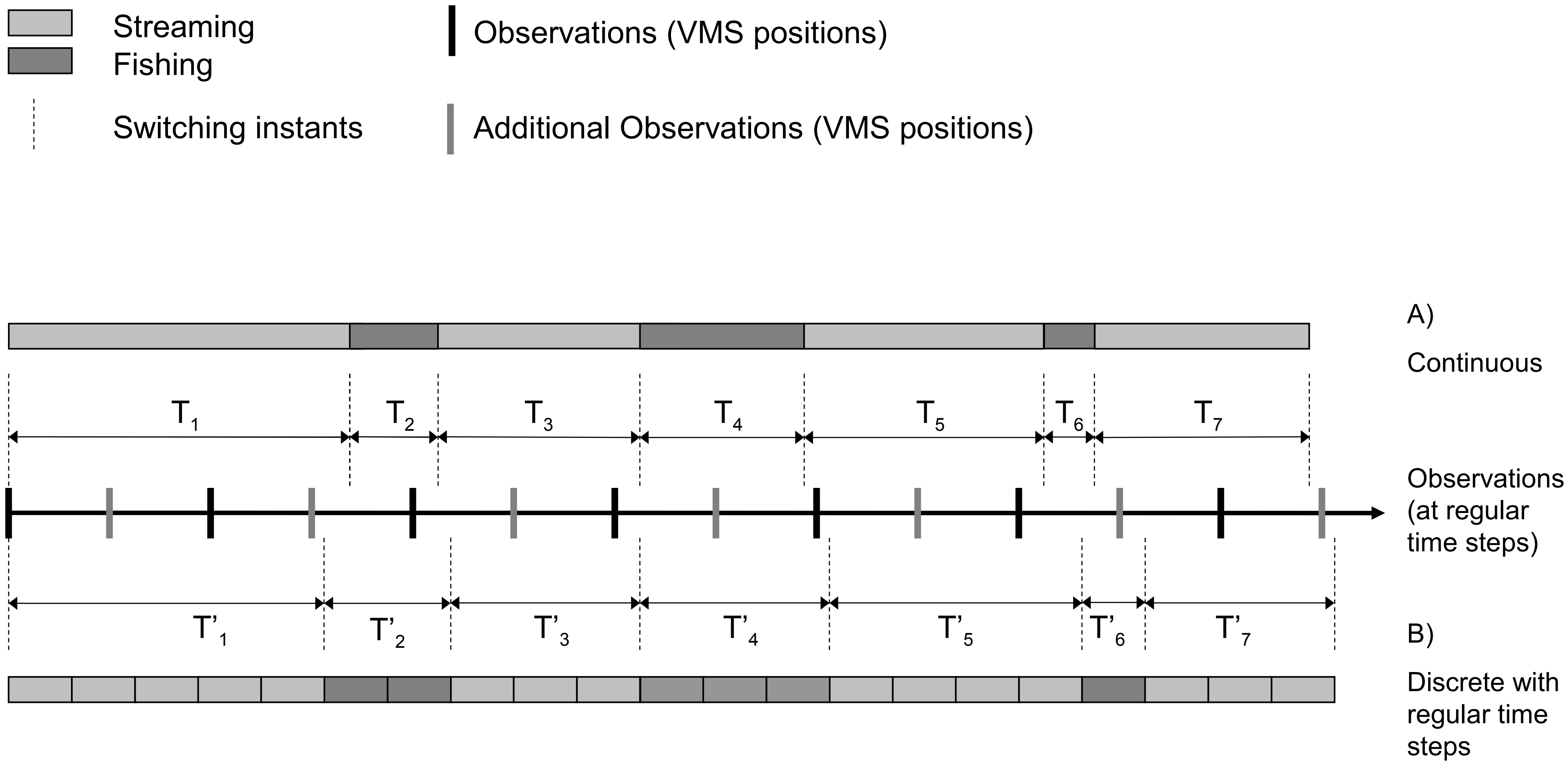
Table 2. Parameters for the simulated pathways and various scenarios. Speeds are drawn at random in Normal distributions with indicated mean and sd. Turning angles are drawn in Wrapped-Cauchy distributions with indicated concentration parameters. Symbol ‘‘ indicates that the characteristics are identical to those of the scenario 1. Symbol ‘‘‘‘ indicates that the characteristics are identical to those of the scenario 8

Scenario	Speeds				Turing angles		Lag-Time	Missing- values	Frequency of observations
	Mean		sd		Concentration parameter of W-Cauchy				
	Fishing	Steaming	Fishing	Steaming	Steaming	Fishing			
1	4	10	1.5	1.5	0.5	0.2	N	N	1
2	4	10	0.1	0.1	0.9	0.5	N	N	1
3	4	10	1.5	1.5	0.5	0.2	Constant =0.1	N	1
4	4	10	1.5	1.5	0.5	0.2	Constant =0.5	N	1
5	4	10	1.5	1.5	0.5	0.2	Constant =0.9	N	1
6	4	10	1.5	1.5	0.5	0.2	Variable~U(0,0.1)	N	1
7	4	10	1.5	1.5	0.5	0.2	Variable~U(0,1)	N	1
8	4	10	1.5	1.5	0.5	0.2	Variable~U(0,0.5)	N	1
9	4	10	1.5	1.5	0.5	0.2	Variable~U(0,0.5)	5%	1
10	4	10	1.5	1.5	0.5	0.2	Variable~U(0,0.5)	10%	1
11	4	10	1.5	1.5	0.5	0.2	Variable~U(0,0.5)	20%	1
12	4	10	1.5	1.5	0.5	0.2	Variable~U(0,1)	N	3

Table 3. Performance of the classification of the behavioural states using proportion [0,1] of simulated behavioural (1 = “Steaming”, 2 = “Fishing”, 3 = “Stopping”) that were correctly (in bold) or wrongly allocated.

	Simulated behavioural state	1			2			3		
	allocated behavioural state	1	2	3	1	2	3	1	2	3
Scenario	1	**0,92**	0,07	0,02	0,08	**0,92**	0,00	0,29	0,14	**0,57**
	2	**0,92**	0,07	0,02	0,08	**0,92**	0,00	0,14	0,14	**0,71**
	3	**0,98**	0,02	0,00	0,08	**0,92**	0,00	0,00	0,00	**1,00**
	4	**0,54**	0,46	0,00	0,31	**0,69**	0,00	0,43	0,57	**0,00**
	5	**0,98**	0,02	0,00	0,08	**0,92**	0,00	0,00	0,00	**1,00**
	6	**0,98**	0,02	0,00	0,08	**0,92**	0,00	0,00	0,00	**1,00**
	7	**0,81**	0,14	0,05	0,23	**0,69**	0,08	0,14	0,43	**0,43**
	8	**0,98**	0,02	0,00	0,08	**0,92**	0,00	0,00	0,14	**0,86**
	9	**0,81**	0,14	0,05	0,23	**0,69**	0,08	0,00	0,57	**0,43**
	10	**0,86**	0,12	0,02	0,08	**0,92**	0,00	0,00	0,43	**0,57**
	11	**0,85**	0,14	0,02	0,08	**0,92**	0,00	0,14	0,29	**0,57**
	12	**0,98**	0,02	0,00	0,08	**0,92**	0,00	0,00	0,00	**1,00**

Figure1



Streaming
Fishing
Switching instants
Observations (VMS positions)
Additional Observations (VMS positions)
A)
Continuous
T1
T2
T3
T4
T5
T6
T7
Observations (at regular time steps)
T'1
T'2
T'3
T'4
T'5
T'6
T'7
B)
Discrete with regular time steps

Figure2

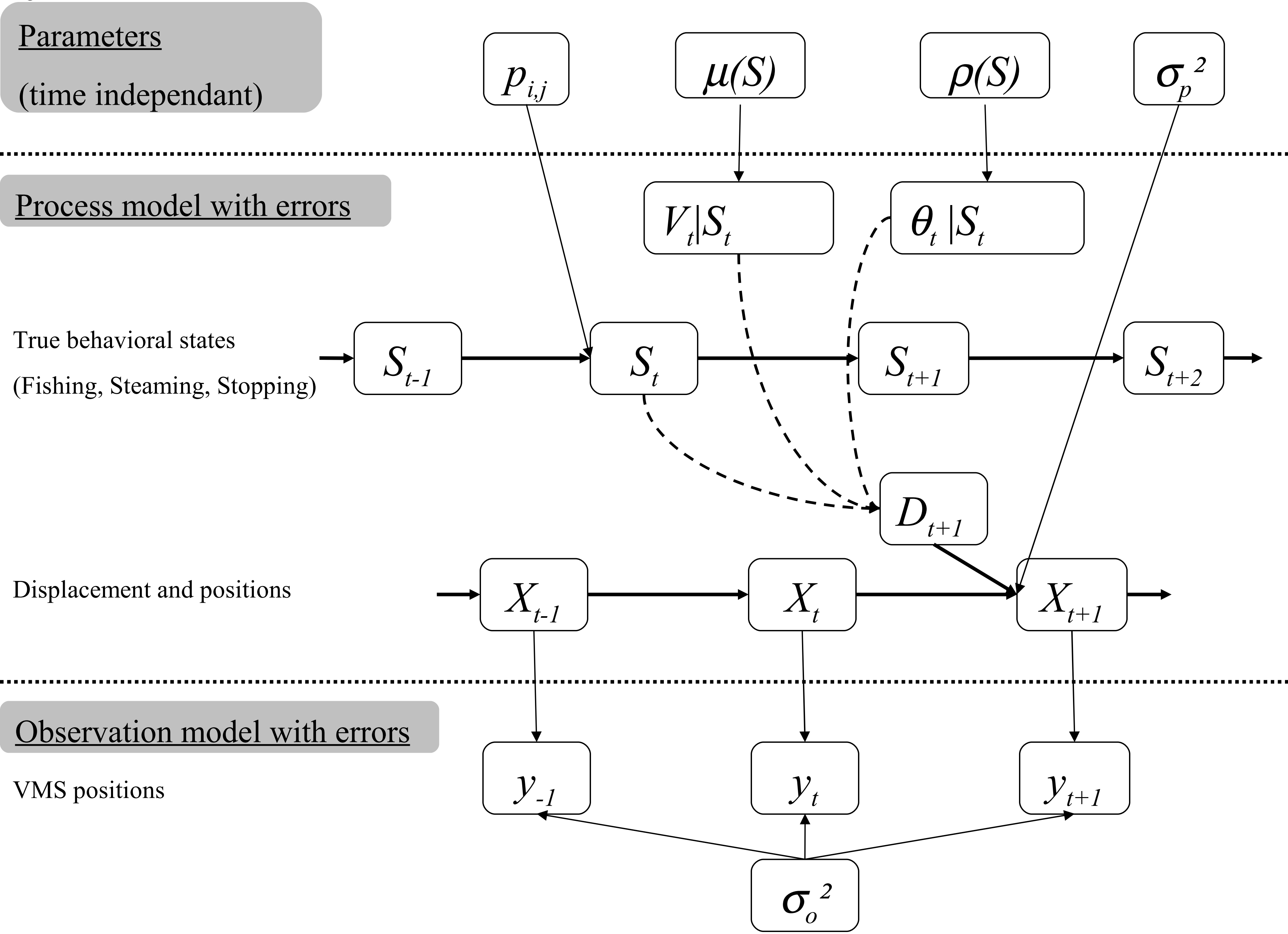

Parameters
(time independant)
$p_{i,j}$
$\mu(S)$
$\rho(S)$
σ_p^2
Process model with errors
$V_t|S_t$
$\theta_t|S_t$
True behavioral states
(Fishing, Steaming, Stopping)
S_{t-1}
S_t
S_{t+1}
S_{t+2}
D_{t+1}
Displacement and positions
X_{t-1}
X_t
X_{t+1}
Observation model with errors
VMS positions
y_{-1}
y_t
y_{t+1}
σ_o^2

Figure3

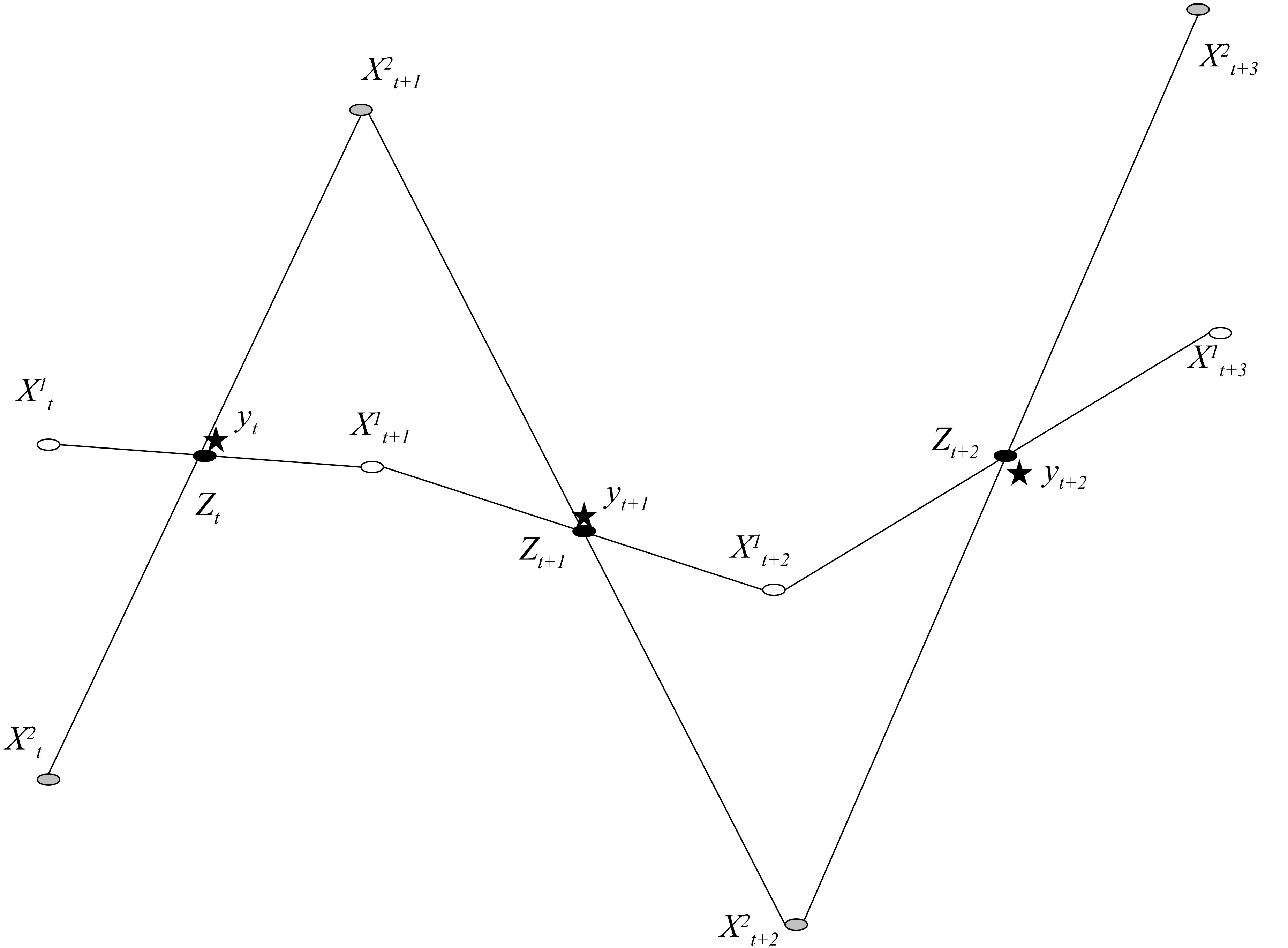


Figure4

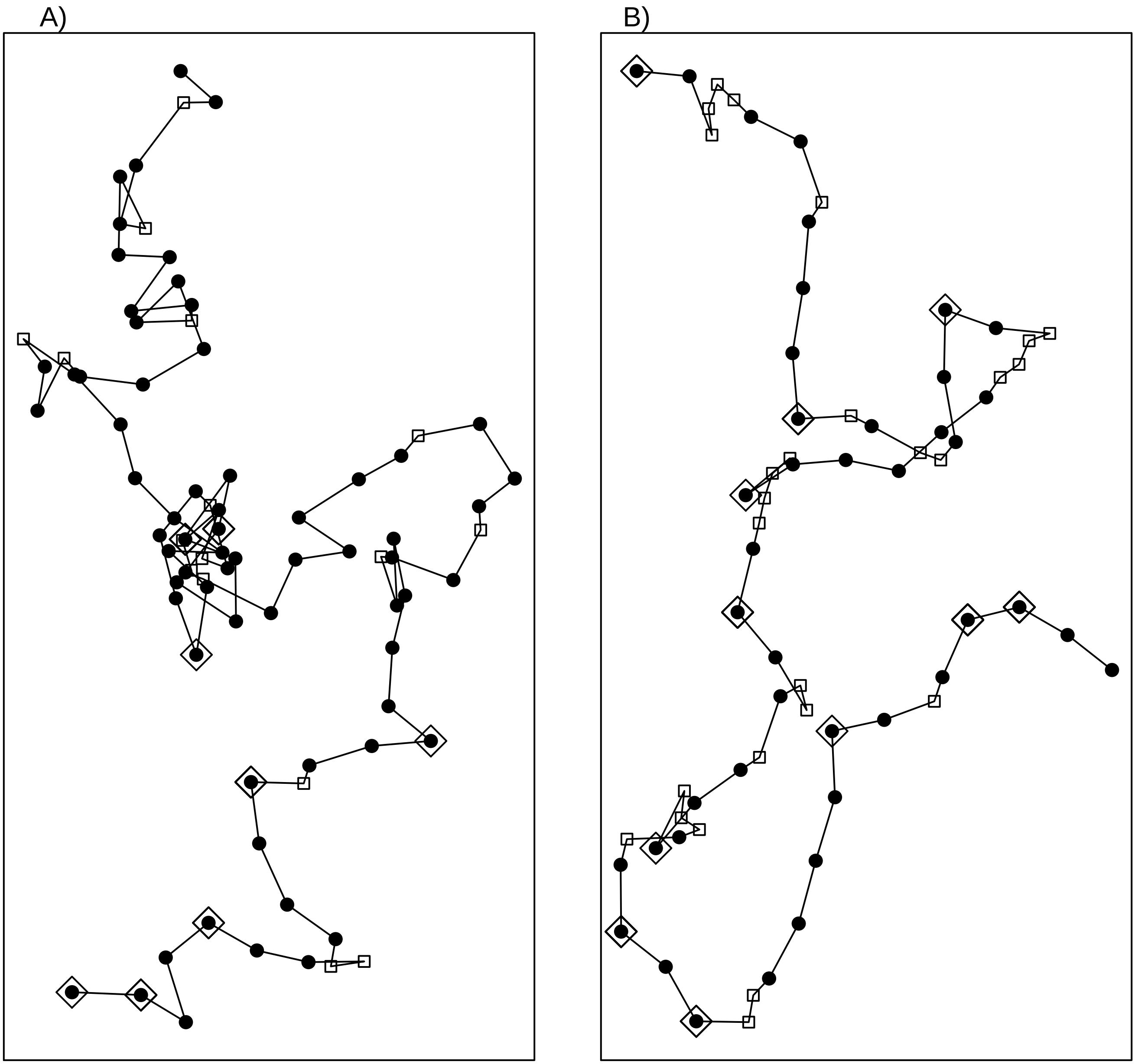


Figure5

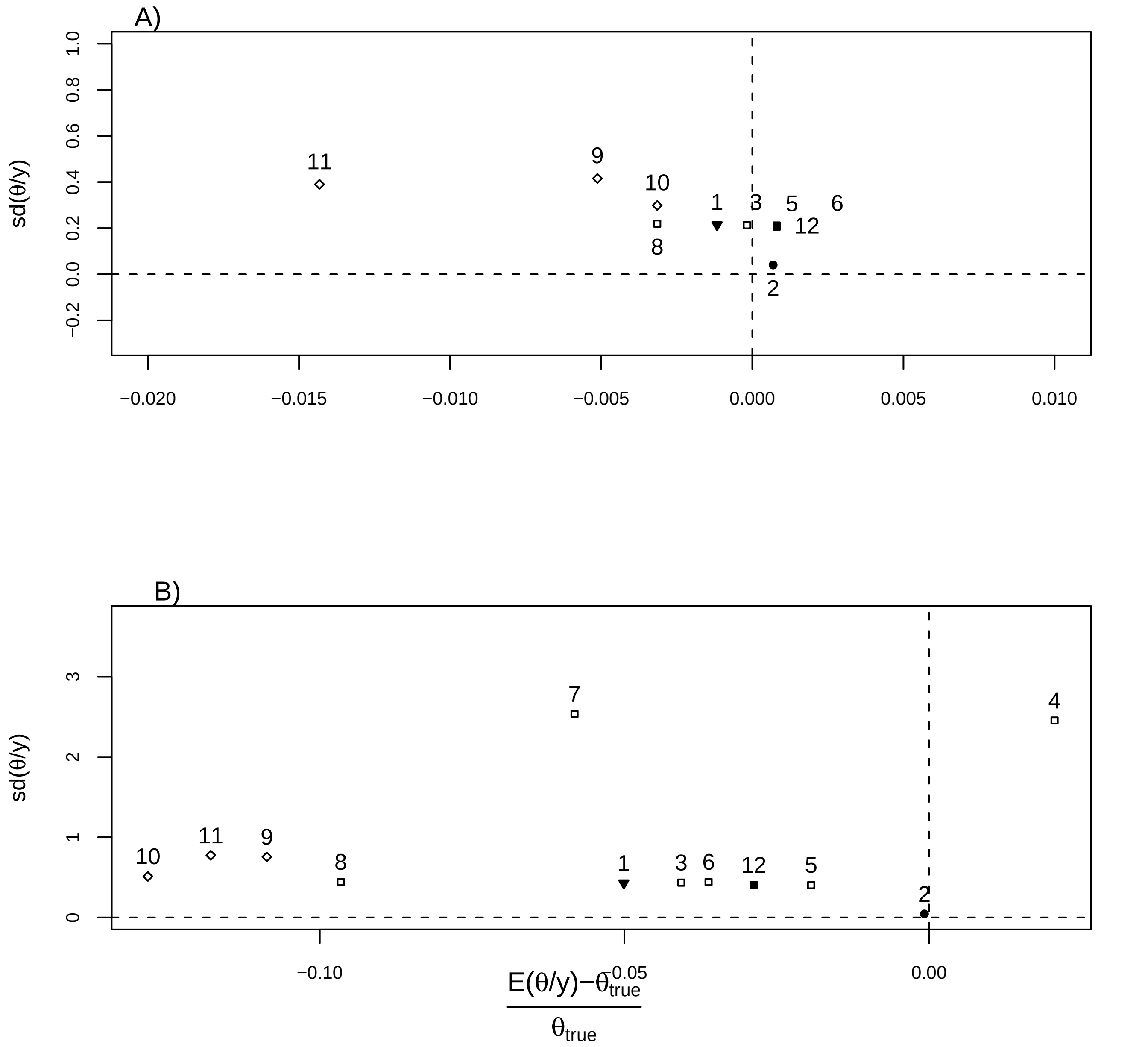

A)
sd(θ/y)
1.0
0.8
0.6
0.4
0.2
0.0
−0.2
11
9
10
1
3
5
6
12
8
2
−0.020
−0.015
−0.010
−0.005
0.000
0.005
0.010
B)
sd(θ/y)
3
2
1
0
7
4
10
11
9
8
1
3
6
12
5
2
−0.10
−0.05
0.00
E(θ/y)−θtrue
θtrue

Figure6

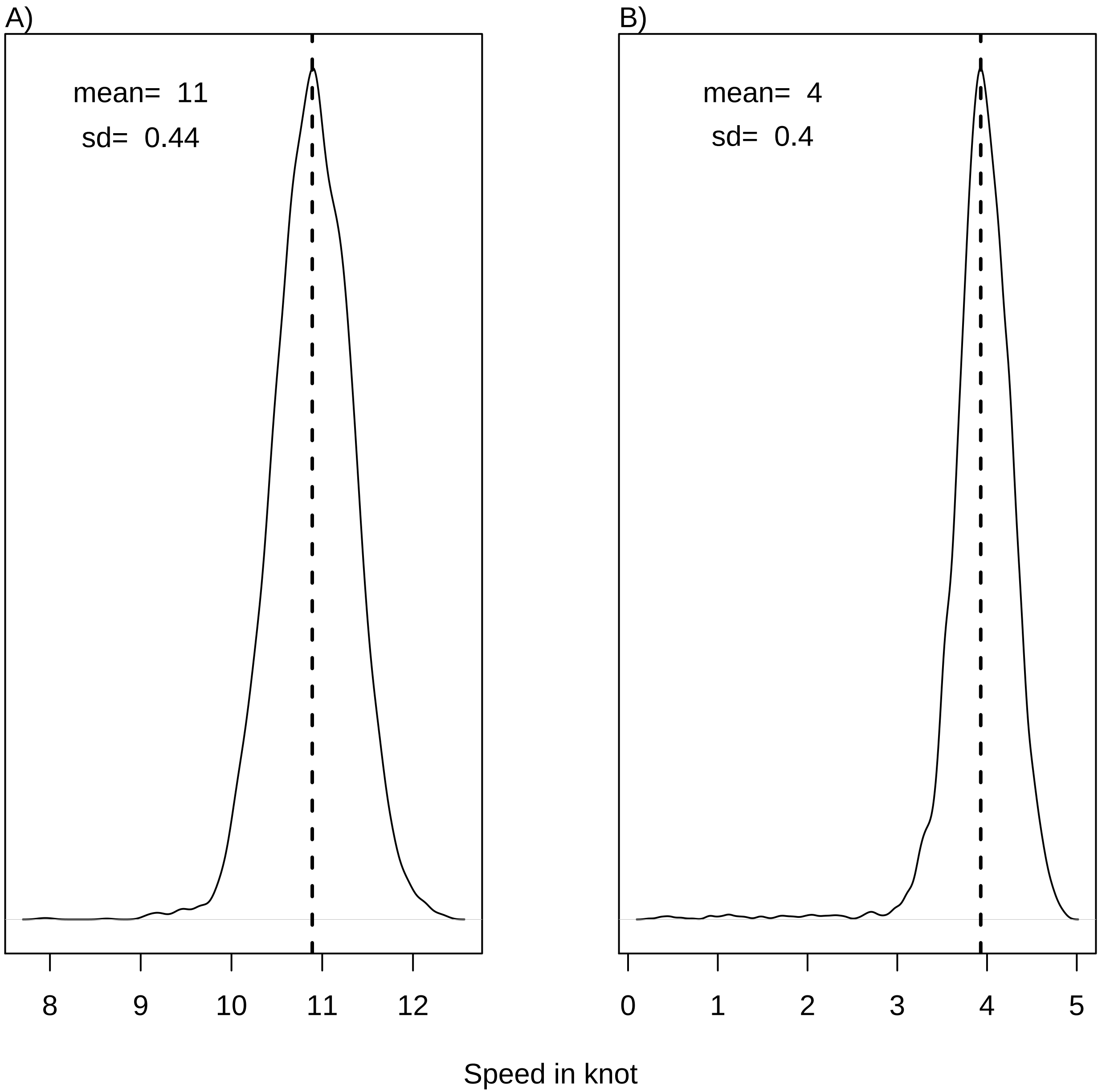
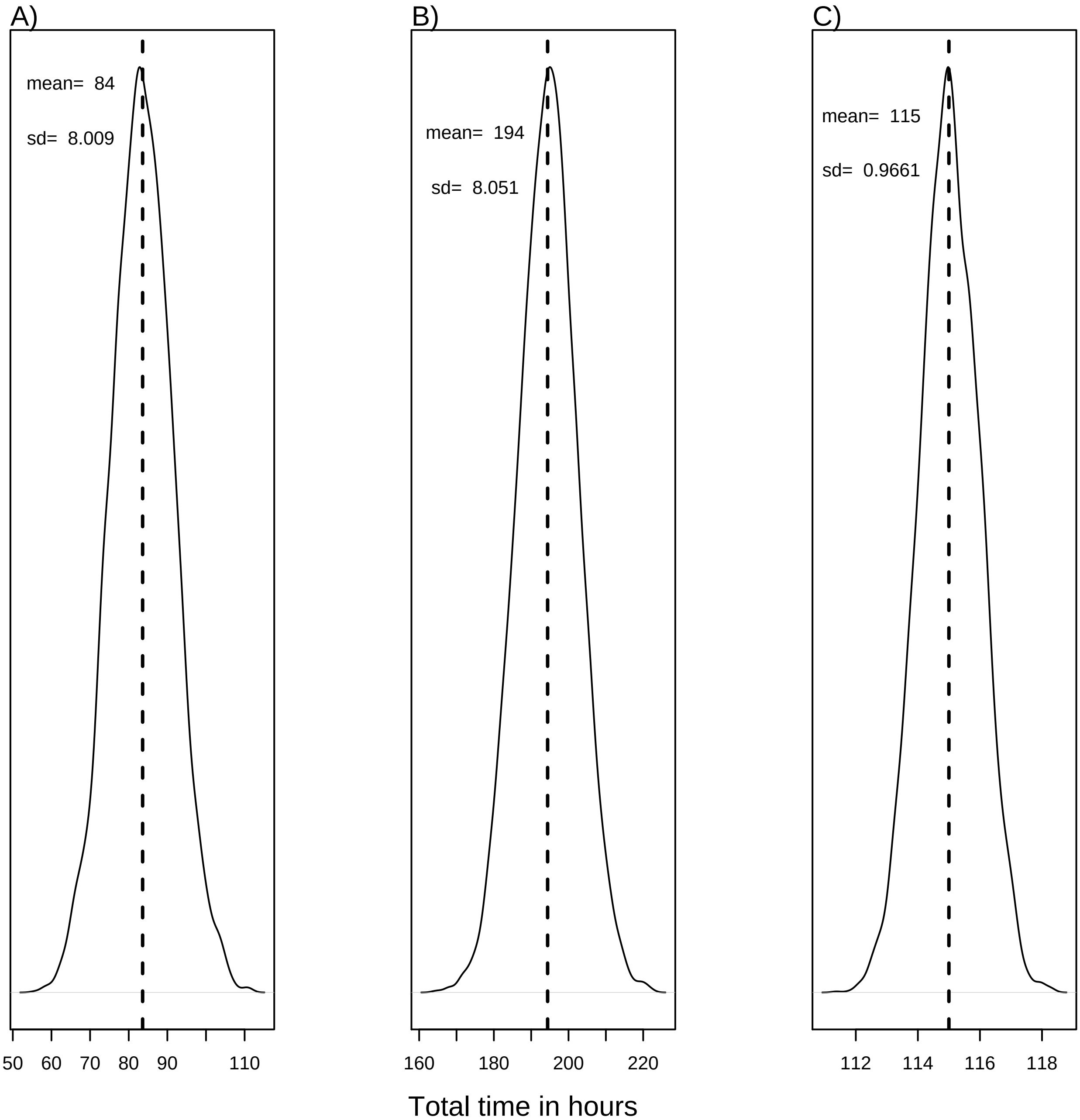
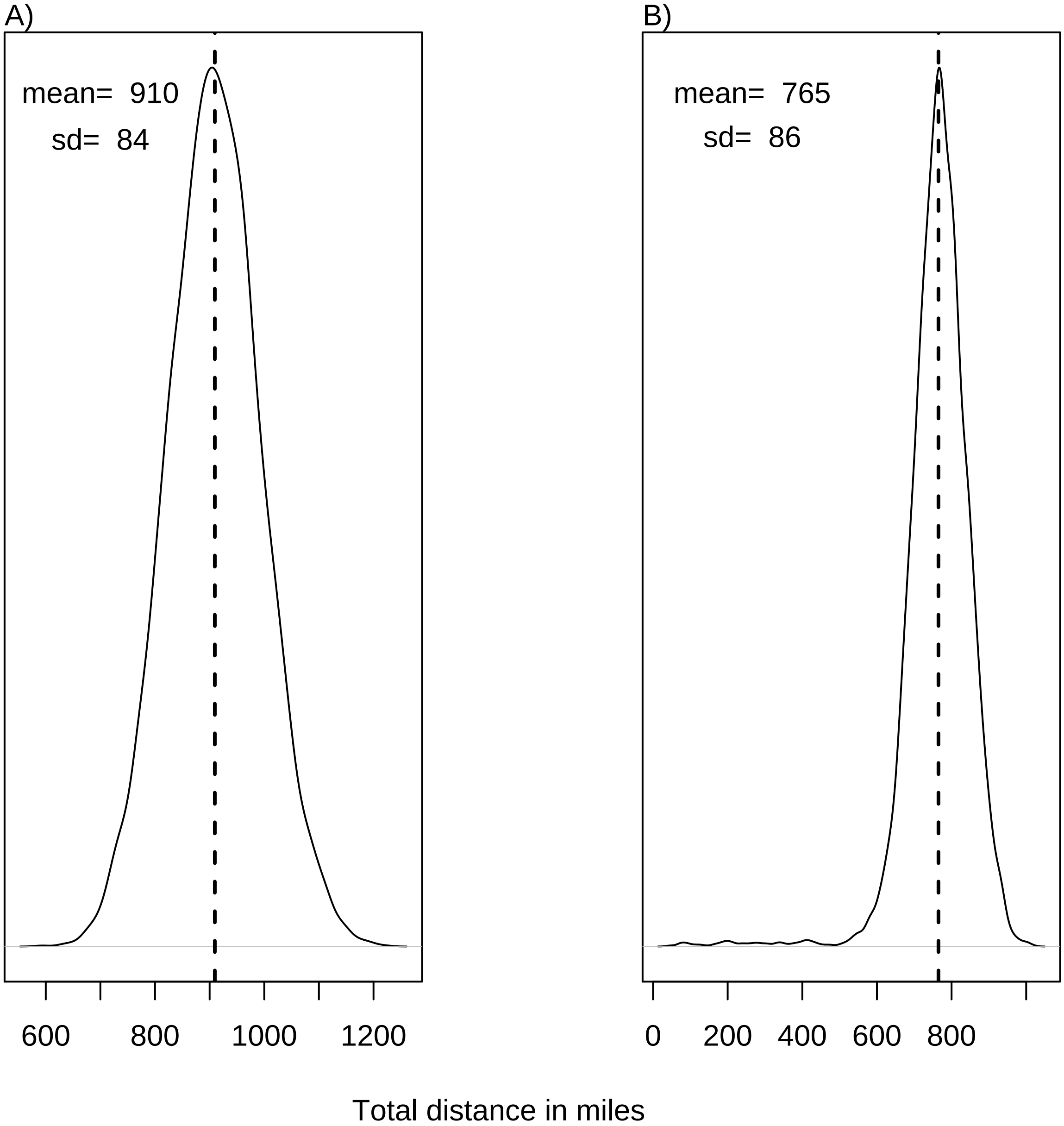


Figure7



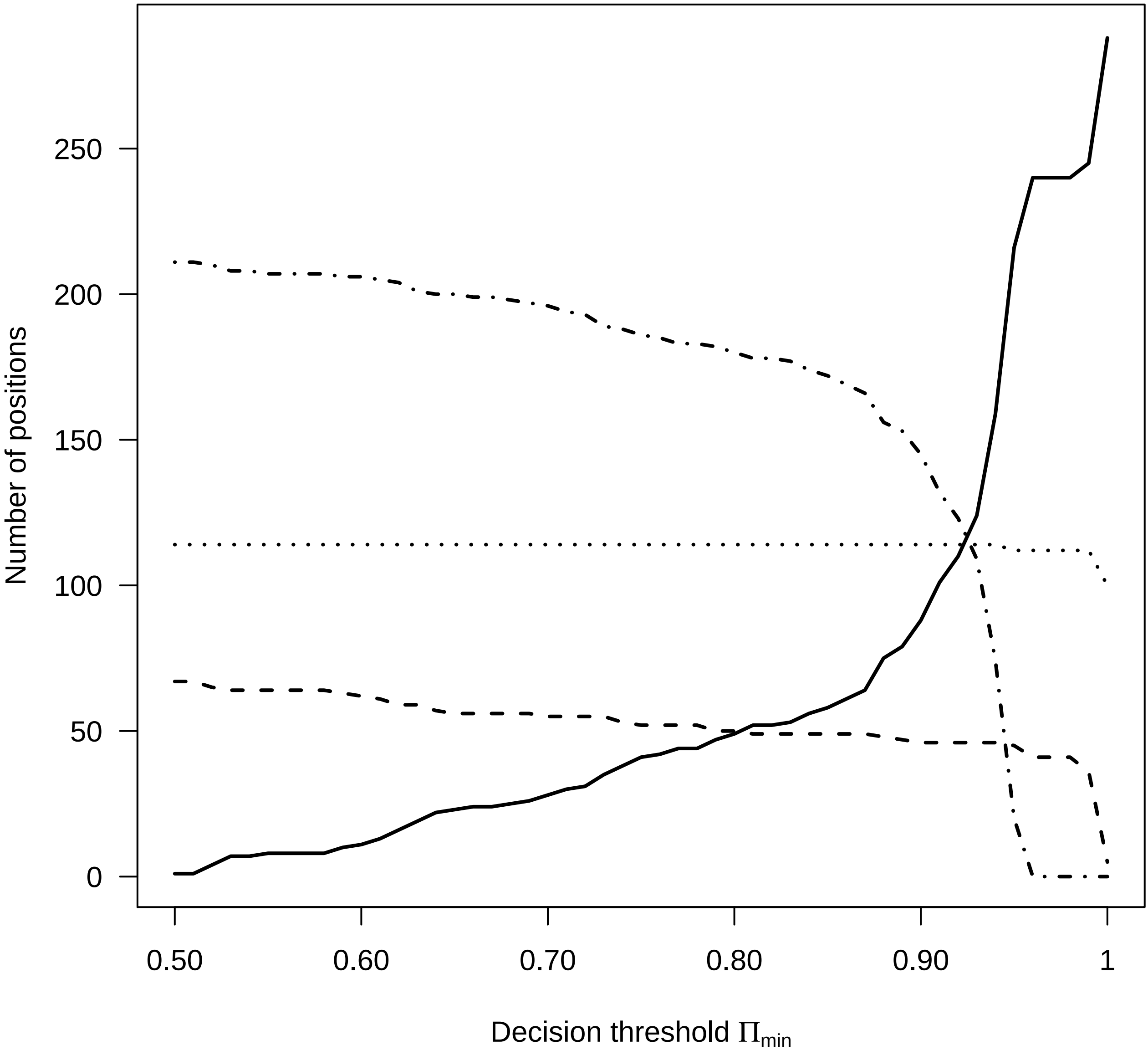
A)
mean= 84
sd= 8.009
50
60
70
80
90
110
B)
mean= 194
sd= 8.051
160
180
200
220
C)
mean= 115
sd= 0.9661
112
114
116
118
Total time in hours

Figure8



A)
mean= 910
sd= 84
600
800
1000
1200
B)
mean= 765
sd= 86
0
200
400
600
800
Total distance in miles

Figure9



Number of positions
250
200
150
100
50
0
0.50
0.60
0.70
0.80
0.90
1
Decision threshold Π_{min}

Figure Captions

Figure 1. Switching time and observation process in continuous time (Panel a) and discrete (Panel b).

Figure 2. Directed acyclic graph for the hierarchical model with hidden Markov chain for behavioral states (see text for the definition of the parameters and variables)

Figure 3. Example of two different true paths $\{X^1{}_t\}$ and $\{X^2{}_t\}$ leading to the same interpolated positions $\{Z_t\}$ (time-lag= 0.5). The paths show sequence of two dimensional positions in a arbitrary Cartesian coordinate system.

Figure 4. Simulated pathways and associated behavioral state ("Steaming" = solid circle, "Fishing" = square, "Stops" = diamond) for scenarios 1 (A) and 2 (B) (see Table 3 for the definition of the scenarios).

Figure 5. Performance of the estimation method for the mean speed associated to the behavioral states "Steaming" (A) and "Fishing" (B) for each scenario 1-12. *x*-axis : relative discrepancy between the estimated and the simulated mean. *y*-axis : Bayesian uncertainty measured as the standard deviation of the posterior distribution of the mean speed. The scenarios (defined in Table 2) are identified by their number. Panel A: Scenarios 4 and 7 are out of the range of the graph (very high bias and uncertainty). Scenario 4: rel. bias = 0.05 and sd = 3.7 ; Scenario 7: rel. bias = -0.15 and sd = 3.2.

Figure 6. Inferences derived from a real data set. Posterior distribution of speeds while "Steaming" (A) and "Fishing" (B).

Figure 7. Inferences derived from the real data set. Posterior distribution of time spent at "Steaming" (A), "Fishing" (B) and "Stopping" (C).

Figure 8. Inferences derived from the real data set. Posterior distribution of distance traveled while "Steaming" (A) and "Fishing" (B).

Figure 9. Number of time steps (over a total of 398 time steps) identified a posteriori in each behavioral state depending on the decision threshold. Solid line corresponds to "unknown state", dashed line to "Steaming", dotted line to "Stopping" and dot-dashed line to "Fishing".