entropy

MDPI

Article

A Simulation-Based Study on Bayesian Estimators for the Skew Brownian Motion

Manuel Barahona [1,2], Laura Rifo [1,*], Maritza Sepúlveda [3] and Soledad Torres [4]

1 Institute of Mathematics and Statistics, University of Campinas, Rua Sergio Buarque de Holanda 651, 13083-859 Campinas, Brazil; mpereira@ubiobio.cl
2 Facultad de Ciencias, Universidad del Bío-Bío, Av Collao 1202, casilla 5C Concepción, Chile
3 Centro de Investigación y Gestión de Recursos Naturales (CIGREN), Instituto de Biología, Universidad de Valparaíso, Av. Gran Bretaña 1111, 2360102 Valparaíso, Chile; maritza.sepulveda@uv.cl
4 Centre for Research and Modeling of Random Phenomena—Valparaíso (CIMFAV), Facultad de Ingeniería, Universidad de Valparaíso, General Cruz 222, 2362905 Valparaíso, Chile; soledad.torres@uv.cl
* Correspondence: laurarifo@ime.unicamp.br; Tel.: +55-19-3521-6052; Fax: +55-19-3289-5766

Academic Editors: Julio Stern and Adriano Polpo
Received: 22 May 2016; Accepted: 22 June 2016; Published: 28 June 2016

Abstract: In analyzing a temporal data set from a continuous variable, diffusion processes can be suitable under certain conditions, depending on the distribution of increments. We are interested in processes where a semi-permeable barrier splits the state space, producing a skewed diffusion that can have different rates on each side. In this work, the asymptotic behavior of some Bayesian inferences for this class of processes is discussed and validated through simulations. As an application, we model the location of South American sea lions (*Otaria flavescens*) on the coast of Calbuco, southern Chile, which can be used to understand how the foraging behavior of apex predators varies temporally and spatially.

Keywords: skew Brownian motion; Bayes estimator; consistency; *e*-value

1. Introduction

For some problems involving time series from continuous variates, considering semi-permeable barriers allows us to construct more accurate models for the phenomenon studied. For instance, in finance we could consider psychological barriers involving asset prices which lead to different decision making procedures on each side of the barrier.

We propose here a model capturing that feature in terms of local times, giving rise to the so-called skew Brownian motion (sBm).

The sBm has attracted interest within other facts due to its relationship with diffusions with discontinuous coefficients or related to media with semi-permeable barriers. It states the first example of the solution of a stochastic differential equation having the local time of the solution as the drift [1].

Recently, several papers have considered the sBm in modeling or in simulation issues, as well as in some optimization problems. See the review by Lejay [2] for references on the subject, and for a survey on the various possible constructions and applications of the sBm.

In this paper, we are interested in the statistical estimation of the skewness parameter when we observe a trajectory of the process at an equally spaced time grid.

From the statistical point of view, this problem represents an intermediate analysis between the classical problem of drift estimation in a diffusion [3,4], and the estimation of the variance of a diffusion, see [5,6] and references therein, where the probability measures generated by the trajectories are singular for different values of the parameter.

As far as we know, there are few works about the estimation of this parameter. In [7], the authors assume that the sBm is reflected at levels 0 and 1 to ensure ergodicity, and in [8], the maximum likelihood estimator is constructed and its asymptotic behavior analyzed, under the null hypothesis of the sBm being the standard Brownian motion. Under general hypotheses, however, the problem is not analytically handled and has not been studied yet.

In this paper, we present a simulation-based analysis on the asymptotic behavior of the Bayes estimators for the skewness parameter, showing that their consistency is supported. The methodology is applied to the trajectory that South American sea lions (SASL, *Otaria flavescens*) follow during their foraging trips, understood as the feeding trips at sea alternating with visits to colonies in order to rest and for lactating females to nurse pups ([9]), on the coast of southern Chile.

The paper is organized as follows. Section 2 provides the model settings and the basic definitions. Section 3 presents a simulation-based analysis on the asymptotic behavior of the mean, mode and quantiles of the posterior distribution. Numerical approximations and real data analysis are exemplified in Section 4. Finally, in Section 5, some aspects of the results obtained are discussed.

2. Model Formulation

2.1. Skew Brownian Motion

Consider the following stochastic differential equation (SDE), for a process $X \in \mathbb{R}$, of the form

$$X_t = x + \sigma B_t + \theta\, l_t^x, \tag{1}$$

where $x \in \mathbb{R}$, $(B_t)_{t\geq 0}$ is a standard Brownian motion, and l_t is the local time of the unique strong solution X at 0.

Given $T > 0$ fixed, for any interval $I \subset [0, T]$, define

$$l(I, A) := \lambda(\{s \in I; X_s \in A\}), \quad A \subset \mathbb{R},$$

where λ is the Lebesgue measure on $\mathbb{R}^+$, and $l(I, \cdot)$ denotes the occupation measure on $\mathbb{R}$ of X during the time interval I. If $l(I, \cdot)$ is absolutely continuous with respect to the Lebesgue measure λ on $\mathbb{R}$, we denote its Radon–Nikodym derivative by

$$l(I, x) = \frac{dl(I, \cdot)}{d\lambda},$$

and we call $l(I, x)$ the occupation density of X at level x during the time interval I. We write $l_t(A) = l([0, t], A)$ and $l_t(x) = l([0, t], x)$ for the occupation measure and for the density (local time), respectively.

An equivalent alternative for the definition of the local time $l^x = \{l_t^x : 0 \leq t \leq +\infty\}$ at level zero of the (unknown) solution X of the Equation (1), departing from x, is given by the following limit,

$$l_t^x = \lim_{\epsilon \to 0} \frac{1}{2\epsilon} \int_0^t 1_{(-\epsilon,\epsilon)}(X_s) ds.$$

Recall that a strong solution X of a stochastic differential equation driven by (1) on a given probability space $(\Omega, \mathcal{F}, \mathbb{P})$, with respect to the fixed Brownian motion B and independent initial condition x over this probability space, is a stochastic process $(X_t, t \geq 0)$ satisfying:

1. X is adapted to the filtration $(\mathcal{F}_t)$, where $\mathcal{F}_t := \sigma(B_s, 0 \leq s \leq t)$;
2. X is a continuous process;
3. $\mathbb{P}(X_0 = x) = 1$;
4. with probability one, we have $X_t = x + \sigma B_t + \theta\, l_t^x, \forall t \geq 0$.

Then, the skew Brownian motion (sBm), $X = \{X_t : 0 \leq t \leq +\infty\}$, can be defined as the strong solution of the stochastic differential Equation (1), where $B = \{B_t : 0 \leq t \leq +\infty\}$ is a standard Brownian motion defined on a given probability space.

For the study of the strong uniqueness of the solution of stochastic differential equations and local times (see [10]), and [2] for a comprehensive survey on the sBm and the SDE involving the local time.

In Equation (1), the initial condition is $x \geq 0$ (the case $x < 0$ is analogous), σ is associated to the volatility parameter, and $\theta \in [-1, 1]$ is the *skewness* parameter.

2.2. Transition Probabilities

Consider the SBm X with parameter $\theta \in [-1, 1]$ defined in Equation (1), for $t \in [0, T]$, and the sampling scheme denoted by $X_i := X_{i\Delta}$, $i = 0, \ldots, n$, and $\Delta = T/n$. As shown in [11], the transition density of X, shown in Figure 1, is given by

$$q(t, x, y \mid \theta, \sigma^2) = \phi_{t\sigma^2}(y - x) + sgn(y)\theta\phi_{t\sigma^2}(|x| + |y|), \tag{2}$$

where

$$\phi_{t\sigma^2}(x) = \frac{1}{\sqrt{2\pi t}\sigma} \exp\Big(-\frac{x^2}{2t\sigma^2}\Big)$$

is the density of a Gaussian random variable with variance $t\sigma^2$ and mean 0, and $sgn(x) = 1$ if $x > 0$, $sgn(x) = -1$ if $x < 0$ and $sgn(x) = 0$ if $x = 0$.

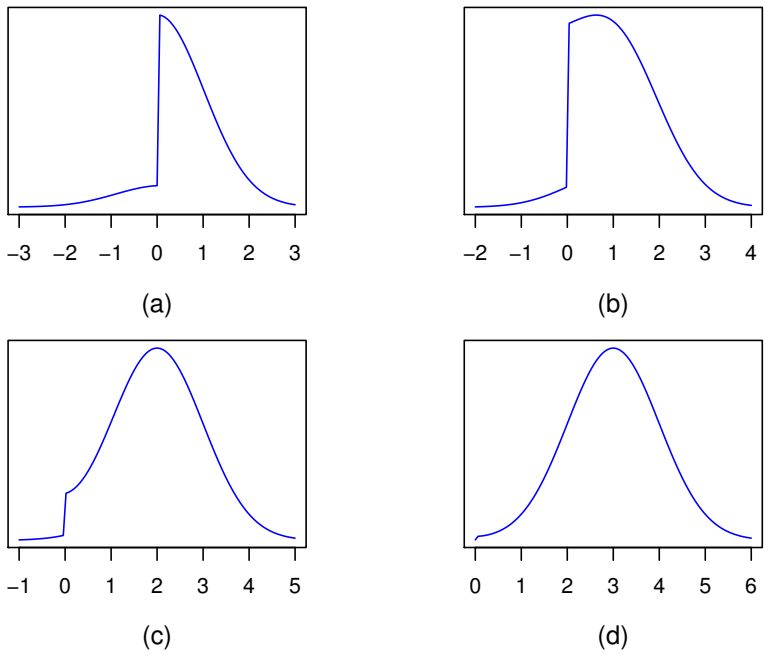

Figure 1. Transition kernel $q(t, x, y \mid \theta, \sigma^2)$, with $t = 1$, $\sigma^2 = 1$, considering $\theta = 0.8$ and: (**a**) $x = 0.1$; (**b**) $x = 1$; (**c**) $x = 2$; (**d**) $x = 2.5$.

We can rewrite the transitions between states as

$$q(t, x, y \mid \theta, \sigma^2) = \begin{cases} \phi_{t\sigma^2}(y - x) + \theta\,\phi_{t\sigma^2}(y + x) & \text{if } y > 0 \\ \phi_{t\sigma^2}(y - x)\,(1 - \theta) & \text{otherwise} \end{cases}, \tag{3a}$$

when $x > 0$,

$$q(t, x, y \mid \theta, \sigma^2) = \begin{cases} \phi_{t\sigma^2}(y - x) - \theta\,\phi_{t\sigma^2}(y + x) & \text{if } y \leq 0 \\ \phi_{t\sigma^2}(y - x)\,(1 + \theta) & \text{otherwise} \end{cases}, \tag{3b}$$

when $x < 0$, and, when $x = 0$,

$$q(t, 0, y \mid \theta, \sigma^2) = \begin{cases} \frac{1}{2}\,\phi_{t\sigma^2}(y)\,(1-\theta) & \text{if } y \leq 0 \\ \frac{1}{2}\,\phi_{t\sigma^2}(y)\,(1+\theta) & \text{otherwise} \end{cases}. \tag{3c}$$

Observe, from the transition kernel, that the parameter θ has a strong influence on the trajectory when it passes close to the origin. Indeed, for a fixed amount of time t, when x is far enough from the origin, the second term is nearly zero. In this case, the trajectory is practically driven by a Brownian motion, almost without influence from θ. On the other hand, when a trajectory approaches the origin from positive values, say, if $\theta < 0$ the trajectory is pushed across the boundary to negative values, and if $\theta > 0$, it has a higher probability to be reflected and remain with a positive value.

In the extreme case $\theta = 1$, the solution to Equation (1) is the reflected Brownian motion above the origin, and when $\theta = -1$, it is the reflected Brownian motion under the origin as soon as the trajectory becomes negative. The case $\theta = 0$ corresponds to the standard Brownian motion.

2.3. Exiting Times

Let us consider $\tau_x = \inf\{t > 0 : X_t = x\}$, the first hitting time of the point $x \in \mathbb{R}$. Let $J = (a, b)$ be any open interval such that the exiting probability $P_x(\tau_{\{a,b\}} < \infty)$ equals 1, for any $x \in J$. Then there exists a continuous, strictly increasing function $S(x)$ on $\mathbb{R}$, such that

$$P_x(\tau_a < \tau_b) = \frac{S(b) - S(x)}{S(b) - S(a)}, \quad x \in J\,.$$

The function S is called the scale function of the process. For more details, refer to [12], for instance. For the sBm, the scale function is given by

$$S(x) = \begin{cases} \alpha^{-1}x & \text{if } x \geq 0 \\ (1-\alpha)^{-1}x & \text{otherwise} \end{cases},$$

where $\alpha = (1+\theta)/2$, as in [2]. This implies that, if $a = -b < 0 < b$,

$$P_0(\tau_{-b} < \tau_b) = 1 - \alpha = (1-\theta)/2\,.$$

Observe that this value does not depend on b, but only on θ: if $\theta = 0$, $P_0(\tau_{-b} < \tau_b) = 1/2$; if $\theta > 0$, $P_0(\tau_{-b} < \tau_b) < 1/2$; and, if $\theta < 0$, $P_0(\tau_{-b} < \tau_b) > 1/2$, for every $b > 0$. In other words, if $\theta < 0$, once the particle reaches the barrier, it is pushed downwards with a probability $(1-\alpha)$ greater than half; and if $\theta > 0$, the particle is pushed upwards with probability α greater than half.

From the point of view of the inference procedure, given an observed trajectory, for those data close enough to the barrier, the proportions of them at each side of the barrier are linear functions of θ. This property can guide us to define an appropriate prior distribution for this parameter, as we discuss in the next section.

3. Bayesian Inference

3.1. Likelihood and Prior Settings

For the model (1), let us consider the parametric space $\Omega = \{\boldsymbol{\omega} = (\theta, \sigma^2) \in \mathbb{R}^2 : \sigma^2 > 0, -1 \leq \theta \leq 1\}$.

For a given sample $\mathbf{d} = (x_1, \ldots, x_n)$, of the observed process at times $(t_1, \ldots, t_n)$, the likelihood function for $\boldsymbol{\omega} \in \Omega$, after observing $\mathbf{d}$, is given by

$$l_{\mathbf{d}}(\boldsymbol{\omega}) = \prod_{i=1}^{n-1} q(\tau_i, x_i, x_{i+1} \mid \theta, \sigma^2)\,,$$

with q as in (3), and $\tau_i = t_{i+1} - t_i$. To ease the notation, let $a_i = a_i(\sigma^2) = \phi_{\tau_i\sigma^2}(x_{i+1} - x_i)$ and $b_i = b_i(\sigma^2) = \phi_{\tau_i\sigma^2}(x_i + x_{i+1})$.

With this notation, the likelihood of (θ, σ^2) is proportional to

$$l_{\mathbf{d}}(\boldsymbol{\omega}) \propto \left(\prod_i a_i(\sigma^2)\right)(1+\theta)^{n_{-+}}(1-\theta)^{n_{+-}} \prod_{\substack{x_j>0\\ x_{j+1}>0}} \left(1 + \frac{b_j(\sigma^2)}{a_j(\sigma^2)}\,\theta\right) \prod_{\substack{x_j<0\\ x_{j+1}<0}} \left(1 - \frac{b_j(\sigma^2)}{a_j(\sigma^2)}\,\theta\right), \quad (4)$$

where n_{-+} is the number of crossings from the negative to the positive side, n_{+-}, the number of crossings from the positive to the negative side, and $b_j/a_j = \exp\{-2x_jx_{j+1}/(\tau_j\sigma^2)\} < 1$.

Observe that, for each step crossing the barrier, if $x_i \leq 0$ and $x_{i+1} > 0$, then $q(\tau_i, x_i, x_{i+1} \mid \theta, \sigma^2) \propto (1+\theta)\,a_i$. Conversely, if $x_i \geq 0$ and $x_{i+1} < 0$, then $q(\tau_i, x_i, x_{i+1} \mid \theta, \sigma^2) \propto (1-\theta)\,a_i$. When the first happens, $\theta = -1$ becomes a root of the likelihood, and when the latter case happens, $\theta = 1$ becomes a root.

On the other hand, for each two consecutive positive or negative steps, $q(\tau_j, x_j, x_{j+1} \mid \theta, \sigma^2)$, as a linear function in θ, has a root, $\pm a_j/b_j$, outside the interval $[-1, 1]$, with maximum in $\theta = 1$ if $x_j > 0$, and in $\theta = -1$, otherwise. This implies, for instance, that if we observe a positive trajectory without crossings, the maximum likelihood estimate for θ is 1, giving more evidence for a reflected Brownian motion, as already pointed out in [8] (Lemma 1).

Suppose now we have prior information about the skewness and the volatility, and let us denote by $f(\boldsymbol{\omega})$ the prior density for $\boldsymbol{\omega}$ quantifying that information. By Bayes's theorem, the posterior density for $\boldsymbol{\omega}$, given the data $\mathbf{d}$, is determined by the relation $f(\boldsymbol{\omega} \mid \mathbf{d}) \propto f(\boldsymbol{\omega}) l_{\mathbf{d}}(\boldsymbol{\omega})$.

By the previous remarks, if we assume prior symmetry around $\theta = 0$, the posterior symmetry for θ basically depends on the polynomial factors defined by the steps performed at each side of the barrier. If, for instance, a trajectory with at least two crossings has more steps at the positive side than at the negative one, then the posterior distribution is skewed to the left, giving more mass probability to $\theta > 0$.

Also, when a trajectory has no crossings, if $x_i > 0$, for all i, then the posterior distribution is skewed to the left, with a modal value in $\theta = 1$, and, conversely, if $x_i < 0$, for all i, then the posterior distribution is skewed to the right, with a modal value in $\theta = -1$.

In this work, we will consider that, a priori, θ and σ^2 are independent random variables. For the inverse of the variance we adopt a Gamma distribution, $G(\alpha_s, \beta_s)$, mainly because of the first product in the likelihood expression (4), so we can obtain a kind of conjugacy for $1/\sigma^2$. On the other hand, from the interpretation of θ as a linear function of a proportion given in Section 2.3 and again for obtaining a partial conjugacy for θ, we adopt for θ a Beta distribution in $[-1, 1]$ with hyperparameters α_t, β_t. With this, the prior density is

$$f(\theta, \sigma^2) \propto (1+\theta)^{\alpha_t - 1}(1-\theta)^{\beta_t - 1} \left(\frac{1}{\sigma^2}\right)^{\alpha_s + 1} \exp\left(-\frac{\beta_s}{\sigma^2}\right) \mathbf{1}_{\Omega}(\theta, \sigma^2)\,,$$

and, consequently, the posterior density can be expressed by

$$f(\theta,\sigma^2 \mid \mathbf{d}) \propto \left(\frac{1}{\sigma^2}\right)^{\alpha_s+1+n/2} \exp\left\{-\frac{1}{\sigma^2}\left(\beta_s+\sum_i \frac{1}{2\tau_i}(x_{i+1}-x_i)^2\right)\right\}(1+\theta)^{n_{-+}}(1-\theta)^{n_{+-}} \times \prod_{\substack{x_j>0\\ x_{j+1}>0}}\left(1+\frac{b_j(\sigma^2)}{a_j(\sigma^2)}\theta\right)\prod_{\substack{x_j<0\\ x_{j+1}<0}}\left(1-\frac{b_j(\sigma^2)}{a_j(\sigma^2)}\theta\right).$$

The last two factors show the dependency between θ and σ^2, given by the updated information after observing $\mathbf{d}$, as stronger the more the trajectory remains near the barrier. As the trajectory moves away from the barrier, those factors converge exponentially to one.

3.2. Hypotheses Testing

One interesting hypothesis for this model is given by the subset $\Omega_0 = \{\boldsymbol{\omega} = (\theta,\sigma^2) \in \Omega : \theta = 0\}$, corresponding to the pure Brownian motion.

Observe that this is a precise hypothesis, that is, it is a sub-manifold of Ω with dimension verifying $\dim(\Omega_0) < \dim(\Omega)$. For any absolutely continuous posterior distribution, the posterior probability of a precise hypothesis is zero. As a test criterion for sharp hypothesis, the Full Bayesian Significance Test (FBST) deals with the posterior probability of a region defined by the hypothesis and the level surfaces of the posterior density $f(\boldsymbol{\omega} \mid \mathbf{d})$.

Given a hypothesis Ω_0, let us define the tangential set, T_0, as

$$T_0 = \{\boldsymbol{\omega} \in \Omega : f(\boldsymbol{\omega} \mid \mathbf{d}) > f_0\}, \quad \text{where } f_0 = \sup_{\Omega_0} f(\boldsymbol{\omega} \mid \mathbf{d}).$$

Intuitively, the tangential set to Ω_0 considers all points more probable than the most probable value inside Ω_0, according to the posterior law.

The *e*-value (briefly denoted by *ev*) for the null hypothesis is an evidence measure defined by

$$ev(\Omega_0) = 1 - \int_{T_0} f(\boldsymbol{\omega} \mid \mathbf{d})d\boldsymbol{\omega}. \tag{5}$$

Therefore, if the tangential set has high posterior probability, the evidence in favor of Ω_0 is small; if it has low posterior probability, the evidence against Ω_0 is small, see [13].

As observed in [14], by the absolute continuity of the posterior distribution on Ω, there is a representation of $f(\boldsymbol{\omega} \mid \mathbf{d})$ such that

$$f_0 = \lim_{\varepsilon\to 0}\left(\sup\{f(\boldsymbol{\omega} \mid \mathbf{d}) : \boldsymbol{\omega} \in V_0(\varepsilon)\}\right), \tag{6}$$

where $V_0(\varepsilon) = \{\boldsymbol{\omega} \in \Omega \mid \text{dist}(\boldsymbol{\omega},\Omega_0) < \varepsilon\}$ is a neighborhood of Ω_0 related to the euclidean distance $\text{dist}(\boldsymbol{\omega},\Omega_0)$. This property allows us to obtain f_0 without having to elicit a prior density conditional on the null hypothesis, using the limit (6) instead.

Under a decision theoretical point of view, the *e*-value defined by (5) allows us to perform a significance test for the hypothesis $H : \boldsymbol{\theta} \in \Omega_0$. Let us consider a decision space

$$\mathcal{D} = \{\text{to reject } H(a_1), \text{not to reject } H(a_0)\},$$

and the loss function $l : \mathcal{D} \times \Omega \to \mathbb{R}$ given by

$$\begin{aligned} l(a_1,\theta) &= w_0(1-\mathbf{1}_{T_0}(\theta)) \\ l(a_0,\theta) &= w_1 + c\mathbf{1}_{T_0}(\theta), \end{aligned}$$

with $w_0, w_1, c > 0$. The value c represents the additional loss of not rejecting H when in fact θ belongs to T_0 and then has posterior density greater than any value in Ω_0. According to [15], the decision rule defined by

$$\text{“reject } H\text{”} \quad \text{if and only if} \quad e-\text{value} < \frac{w_1 + c}{w_0 + c} \tag{7}$$

minimizes the expected value $E(l(a,\theta)|\mathbf{d})$.

We can also measure the evidence for H using the well-known Bayes factor [16]. Let us note that its determination requires us to define a prior distribution on Ω_0, $f(\boldsymbol{\theta}|\Omega_0)$. The sensitivity of Bayes factors to the choice of such priors is one of the usual criticism of that measure (see [17]). Let B denote the Bayes factor against the hypothesis H. According to [18], one might adopt the following decision rule to test H: if $1 < B < 3$ there is little evidence against H; if $3 < B < 20$, the evidence against H is positive; if $20 < B < 150$, it is strong; and if it is greater than 150, the evidence is very strong. These categories are though prescriptive, since there is no explicit mention to any loss function associated to the test.

Finally, we will also compute posterior probabilities, as presented in Section 4, for some regions of interest.

3.3. Considerations on the Asymptotic Behavior of the Bayesian Posterior

This section presents a simulation-based analysis of the asymptotic behavior of the posterior distribution of θ, given σ^2, as the sample size increases.

In [8], a first step towards a convergence result for the maximum likelihood estimator (MLE) has been achieved. The authors characterize the limiting distribution under the null hypothesis of the standard Brownian motion, $\theta = 0$, and the convergence is related to the $\mathcal{G}$-stability (see [6] for the definition and properties). That kind of convergence is proved to be true there in a neighborhood of $\theta = 0$, but even in that case it is not a sufficient condition for the consistency of the estimator. To the best of our knowledge, reference [8] is the only analytical result on the convergence of estimators for the skewness parameter θ of the sBm.

Also, for the MLE, high-frequency estimation, obtained for a sequence of times $\{iT/n\}_i$, is equivalent to long-time estimation, obtained for a sequence of time-steps equal to one. The simulations we performed consider then increasing times $\{1, \ldots, n\}$ and give support for the consistency of the posterior distribution when the sample size n increases, as we can see in what follows.

For this discussion, we present the simulated sampling distribution of the posterior mode, the posterior mean, and the posterior 0.025 and 0.975 quantiles for θ. Those distributions were obtained by a simulation of 10 thousand trajectories with n steps, for $n = 100, 1000$ and $10,000$, for different values of θ ranging from -1 to 1. We consider for the simulations a fixed value of $\sigma^2 = 1$; analogous results were obtained for other values of σ^2.

Figure 2 represents the sampling distribution for the posterior mode of θ, from trajectories generated with $\theta = 0$.

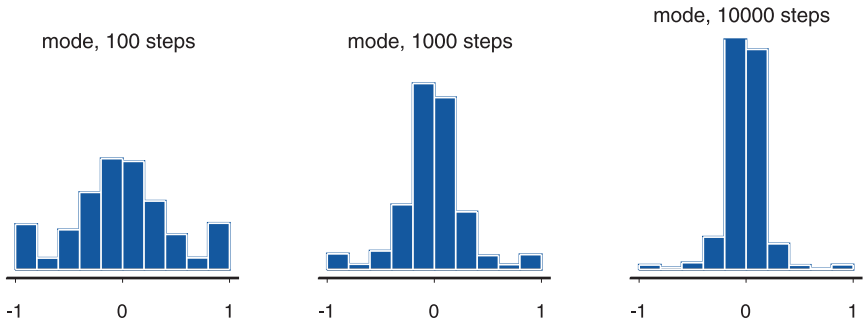

Figure 2. Sampling distribution for the posterior mode given ten thousand trajectories with n steps, n = 100, 1000, 10,000, generated from $\theta = 0$.

Observe that for a trajectory with $n = 100$ steps, the sampling distribution of the posterior mode is trimodal, with modes $-1, 0, 1$. These extreme modes occur because there is a positive probability that a trajectory does not hit the origin, giving then evidences for $\theta = 1$ if the trajectory is mostly positive, and for $\theta = -1$, if it is mostly negative. By symmetry, both cases can happen with the same probability, as confirmed by the simulations. As n increases, this probability tends to zero, as we can see in the histograms for $n = 1000$ and $n = 10,000$ steps, where the mass probability around -1 and 1 almost vanishes.

Figure 3 represents the sampling distribution for the posterior mean and the joint sampling distribution for the posterior quantiles $q(0.025)$ and $q(0.975)$. It is clear that the sampling distribution for the posterior mean concentrates around $\theta = 0$ as n increases. Also, the range $q(0.975)$–$q(0.025)$ diminishes as n increases, showing that the whole posterior distribution concentrates asymptotically.

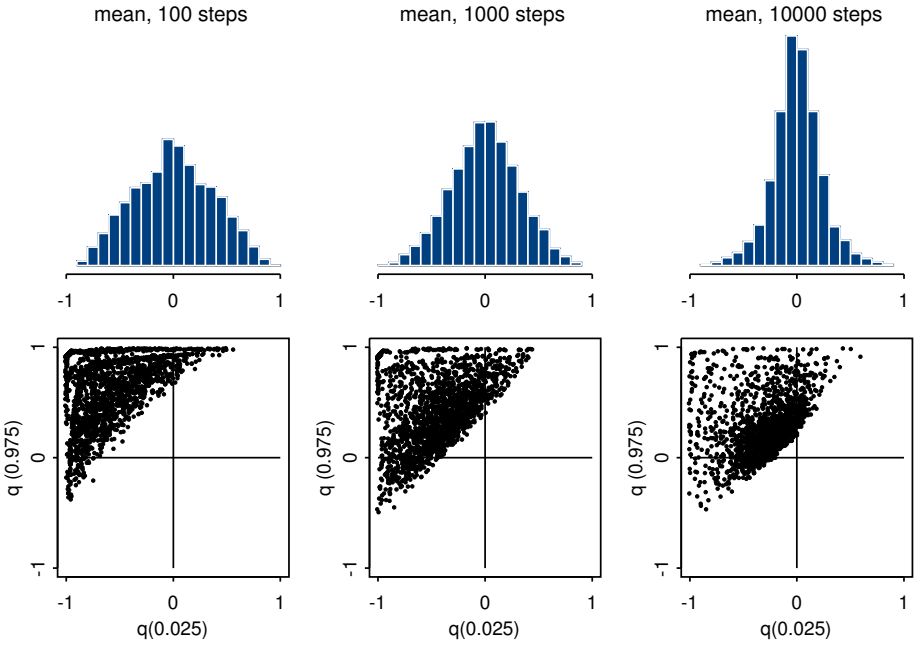

Figure 3. Posterior mean, and quantiles $q(0.025)$ and $q(0.975)$, for trajectories with n steps, n = 100, 1000, 10,000, assuming $\theta = 0$.

The convergence can also be observed in the sampling distribution of the posterior mean for different values of θ, as shown in Figure 4. There, the horizontal axis represents the nominal values for θ, from -1 to 1. For each θ, the solid lines represent the 95% central values of the posterior mean, for a sample of size: (a) $n = 100$, (b) $n = 1000$, (c) $n = 10,000$ steps. Observe that these intervals become sharper as n increases.

We must point out that the number of crossings is a relevant statistic for having a more precise inference on θ. In each graph, the dashed line shows the same interval for the posterior mean, but considering only those trajectories that have five crossings or more, and the dot-dashed line, those with 10 crossings or more. The information gained is remarkable for small samples, as we see in Figure 4, mainly for θ close to 1 or -1.

As seen in Figures 2 and 3, for sample sizes less than 1000, the posterior mode for the skewness parameter presents a more dispersed sampling distribution than the posterior mean. That makes, in some sense, the posterior mean a more reliable estimator for θ. On the other hand, both estimators improve as the number of crossings increases in the sample, and in that sense their consistency is well supported by simulations, for any value of θ.

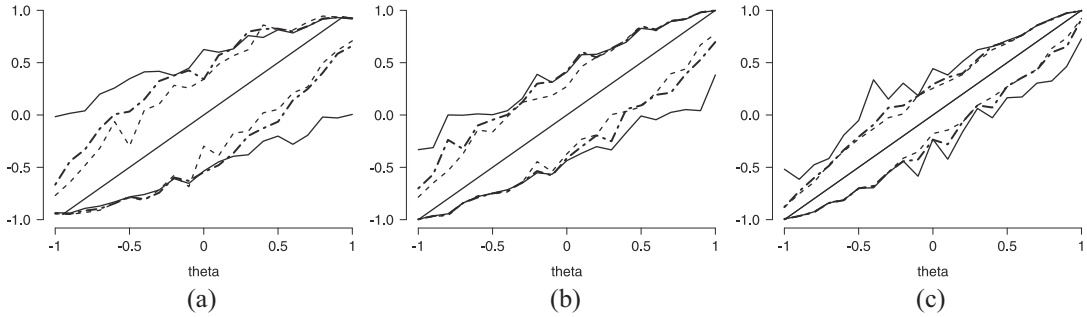

Figure 4. Sampling interval for the posterior mean of θ, with $\sigma^2 = 1$, considering $\theta \in [-1, 1]$ and: (**a**) $n = 100$; (**b**) $n = 1000$; (**c**) $n = 10,000$ steps. For each graphic, the solid line represent the 95% central values of the posterior mean, the dashed line shows the same interval considering only those trajectories that have five crossings or more, and the dot-dashed line, those with 10 crossings or more.

4. Computation for Data

4.1. Simulated Data

In this section we describe the procedure followed to obtain posterior inferences from one single trajectory, and the subsequent point estimates and e-value. For illustrative purposes, let us consider the simulated trajectory of $n = 1000$ steps from a skew-Brownian motion with parameters $\sigma^2 = 1$ and $\theta = 0.8$, represented in Figure 5a.

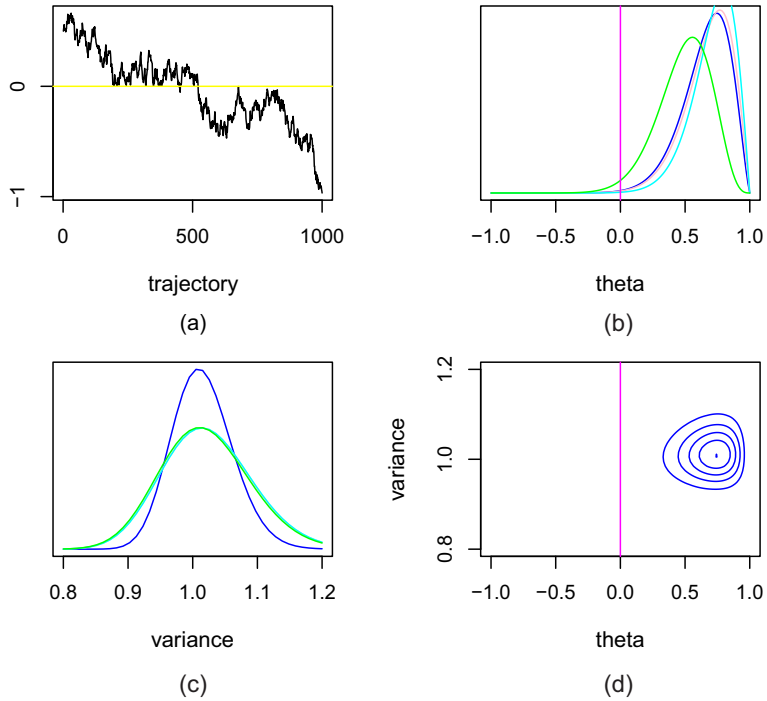

Figure 5. (**a**) Simulated skew-Brownian trajectory, with $n = 1000$ steps, for $\theta = 0.8$ and $\sigma^2 = 1$. Note that the trajectory tends to be more positive than negative; (**b**) Marginal posterior distribution for θ, with prior distribution for θ: Beta(1,1) in blue, Beta(0.5,0.5) in pink, Beta(5,1) in cyan, Beta(1,5) in green. The magenta vertical line represents the hypothesis $\theta = 0$; (**c**) Marginal posterior distribution for σ^2, with prior distribution inverse-Gamma[1,1]; (**d**) Joint posterior distribution for (θ, σ^2), with prior distribution for θ: Beta(1,1) in blue, Beta(0.5,0.5) in pink, Beta(5,1) in cyan, Beta(1,5) in green. The magenta vertical line represents the hypothesis $\theta = 0$.

We have chosen a prior Gamma[1,1] distribution for $1/\sigma^2$ and four different Beta prior distributions for θ, with parameter (α_t, β_t) equals to $(1,1), (0.5,0.5), (5,1), (1,5)$. As the posterior distribution $f(\boldsymbol{\omega} \mid \mathbf{d})$ has a closed analytical form, except for the normalizing constant, we calculated it in a grid with 141×40 points for (θ, σ^2) in $[-1,1] \times [0.8, 1.2]$. The posterior distributions corresponding to each of those priors are represented in Figure 5 in colors blue, pink, cyan and green, respectively.

Let us observe that there is practically no difference among the posterior distributions obtained from the three first priors. This lack of influence is due to the large number of steps close to the barrier, giving, as already mentioned, relevant information about θ. Just when we strongly favor $\theta = -1$, with the Beta(1,5) prior, the difference in the estimates for θ is not unnoticed, as we can see in Table 1. Even in that last case, the e-value in favor of the hypothesis $\theta = 0$ is insignificant, agreeing also with the Bayes factor.

Table 1. Posterior estimates for (θ, σ^2), for different priors on θ, for the simulated trajectory with nominal values $\theta = 0.8$ and $\sigma^2 = 1$. The last two columns show the evidence in favor of the hypothesis $\theta = 0$, given by the e-value and the Bayes factor.

Prior Distribution for θ	Posterior Mode for (θ, σ^2)	Posterior Mean for (θ, σ^2)	e-Value for $\theta = 0$	Bayes Factor against $\theta = 0$
Beta(1,1)	$(0.74, 1.01)$	$(0.66, 1.02)$	0.0029	16.8
Beta(1,5)	$(0.56, 1.02)$	$(0.49, 1.02)$	0.0224	2.55

4.2. Real Data

Seven South American sea lions, SASL, 1 male and 6 females, were captured in July 2009 and June 2010 on the coast in front of Calbuco (41°48′ S; 73°08′ W), southern Chile. SASL were captured as they approached commercial purse-seine vessels to forage during sardine fishing. Animals were brought on board the vessel and then transported to shore where they were anesthetized. On land, the individuals were instrumented with Sea Mammal Research Unit-Satellite Relay Data Logger (SRDL) GPS tags (University of St. Andrews, Scotland), that were glued to the dorsal pelage. The SRDL-GPS tags collected data on animal position, diving behavior and water temperature. These data were processed on board to calculate a series of derived parameters, and then summarized and transmitted via the ARGOS system.

In this work, we modeled the foraging trip locations of one of these SASL (subject ID 96721), as an example of a skew trajectory. Several studies on different sea lion species have demonstrated a fidelity to foraging areas among individuals (e.g., [17–19]), including SASL on the Atlantic coast ([20,21]). Individual fidelity to foraging areas may allow animals to locate productive foraging grounds on successive trips, which could be a beneficial strategy as it confers the advantage of reducing overall travel costs ([18,20]). Based on this, we hypothesized that the SASL will always travel to the same foraging areas, and thus, within the proposed model, θ should be close to -1.

Most foraging trips of this subject were concentrated between La Sebastiana colony and the inner waters of Chiloe Island (see Figure 6). As shown in Figure 7a, there exists a physical barrier at latitude -41.76, which corresponds to the location of the sea lion colony.

As in the simulated example, we consider four different prior distributions for θ, Beta(α_t, β_t), with (α_t, β_t) equals to $(1,1), (0.5,0.5), (5,1), (1,5)$, represented in Figure 7 in colors blue, pink, cyan and green, respectively. We also standardized the observed trajectory, dividing all terms by the standard deviation of the observed differences, and considered a Gamma(1,1) prior distribution for $1/\sigma^2$. Posterior estimates are given in Table 2.

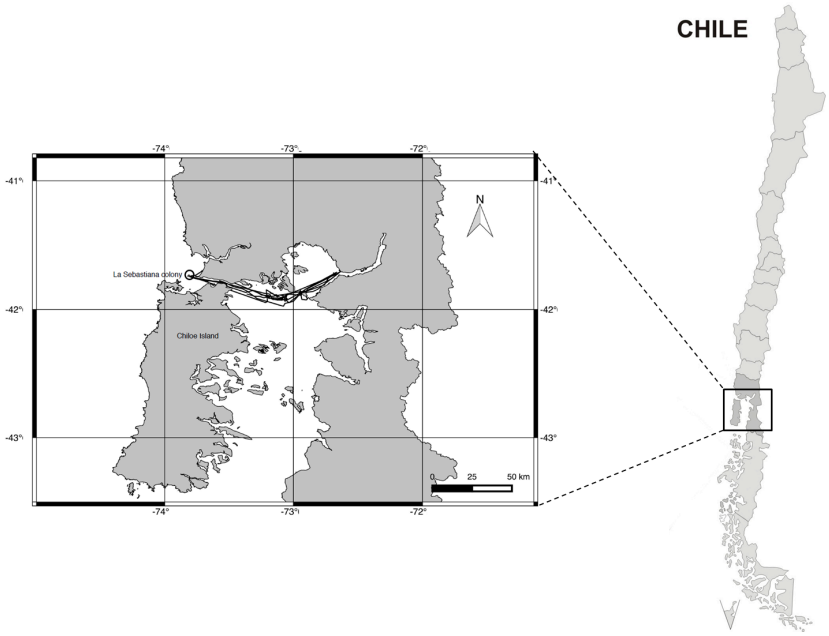

Figure 6. Localization of the foraging area, Calbuco, Chile.

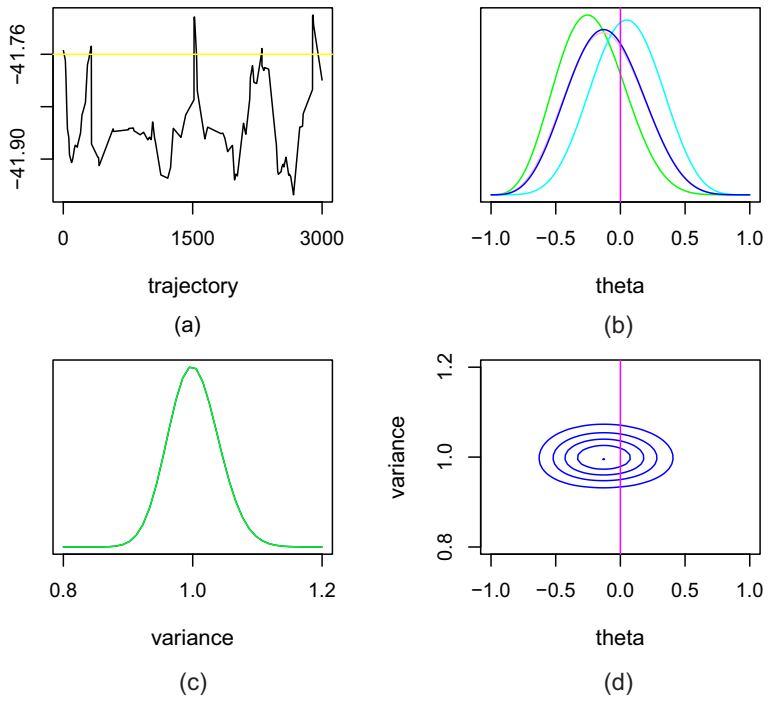

Figure 7. (**a**) Latitude of a sea lion trajectory on the coast in front of Calbuco, southern Chile; (**b**) Marginal posterior density for the skew parameter θ, with prior distribution for θ: Beta(1,1) in blue, Beta(0.5,0.5) in pink, Beta(5,1) in cyan, Beta(1,5) in green. The magenta vertical line represents the hypothesis $\theta = 0$; (**c**) Marginal posterior distribution for σ^2, with prior inverse-Gamma[1,1]; (**d**) Joint posterior distribution for (θ, σ^2), with prior distribution for θ: Beta(1,1) in blue, Beta(0.5,0.5) in pink, Beta(5,1) in cyan, Beta(1,5) in green. The magenta vertical line represents the hypothesis $\theta = 0$.

Table 2. Posterior estimates for (θ, σ^2), for different priors on θ, for the SASL observed trajectory. The last three columns show the evidence in favor of the hypothesis $\theta = 0$, given by the *e*-value, the Bayes factor and the posterior probability of $\theta < 0$.

Prior Distribution for θ	Posterior Mode for (θ, σ^2)	Posterior Mean for (θ, σ^2)	*e*-Value for $\theta = 0$	Bayes Factor against $\theta = 0$	$P(\theta < 0 \mid \mathbf{d})$
Beta(1,1)	$(-0.13, 0.99)$	$(-0.11, 1.00)$	0.656	0.394	0.648
Beta(5,1)	$(0.04, 0.99)$	$(0.04, 1.00)$	0.876	0.283	0.424
Beta(1,5)	$(-0.26, 0.99)$	$(-0.23, 1.00)$	0.355	0.404	0.799

The evidence given by the Bayes factor against H is favorable to this hypotheses. However, considering the decision rule (7) and an uniform prior distribution for θ, we should take the decision to reject H if and only if

$$0.656(w_0 + c) < w_1 + c \iff \frac{2}{3}w_0 < w_1 + \frac{1}{3}c\,, \text{ approximately.}$$

In other words, if we have for instance $w_1 \approx c$, and the cost of the error of type II is at least half the cost of the error of type I, then H is to be rejected.

Also, the probabilities $P(\theta < 0 \mid \mathbf{d}) > 0.4$, for the different priors, indicate that the subject likely did not always travel to the same foraging area but instead some of their trips were to lower latitudes. Biologically, at least three different explanations not mutually exclusive could explain why SASL move to different areas to forage. First, prey aggregations may change location due to variations in oceanographic features, thus obligating predators to change their foraging areas, see [22]. Second, individuals usually segregate and change their foraging locations in order to decrease intra-specific resource competition ([23]). Finally, SASL are considered to be a generalist predator, which means that they consume different prey species [24]. Thus, if the prey abundance and availability in a determined foraging area decrease, the subject may display behavioral plasticity and abandon the area and forage at a different place or look for different prey ([25]).

5. Conclusions

In a model selection problem, some aspects should be taken into account, as theoretical considerations on the question itself or past experience. The observed sample may provide insights about the skew Brownian model, mainly if there are a number of crossings through the barrier, but supplementary considerations are required for choosing the skew Brownian as a proper model for a specific problem.

From the point of view of the estimation, as seen in Figure 3, for sample sizes less than 1000, the posterior mode and, in particular, the maximum likelihood estimator for the skewness parameter present more dispersed sampling distribution than the posterior mean.

That makes, in some sense, the posterior mean a more reliable estimator for θ. On the other hand, both estimators improve as the number of crossings increases in the sample, and in that sense, their consistency is well supported by simulations.

A possible generalization of this problem is to consider that, on each side of the barrier, there are different values for the volatility parameter, σ_1^2 and σ_2^2, allowing in that way different dynamics depending on the barrier. On this extended parametric space, the proposed methodology can be applied as well. Considering that parametrization, we obtain a precise hypothesis stating that those diffusions have the equal value to the parameter $\sigma_1^2 = \sigma_2^2$, whose evidence can be measured by the *e*-value, by instance. Another possible extension consists of considering more than one barrier and several diffusion processes between them, or, even more, considering a time-dependent barrier, underlying dynamic models problems.

The analysis proposed here could be used to model behavioral changes in apex predators, such as marine mammals. Understanding the foraging behavior of the animals, and how such behavior varies temporally and spatially may be an indicator of ecosystem events and links shifts in predator and prey dynamics with environmental variability.

Acknowledgments: The authors would like to thank the reviewers for having provide very helpful corrections and guidance for clarifying some of the discussions. Barahona is a PhD student with partial CAPES Grant at the University of Campinas. Soledad Torres was partially supported by Proyecto Anillo ACT1112, ECOS C15E05, Fondecyt 1130586 and Mathamsud 16MATH03. Financial support for SASL data was provided by INNOVA-CORFO 07CN13IPM-170. All SASL handling and tagging procedures were enabled by Subsecretaria de Pesca Permits N 2799/2008 and 1737/2010, and approved by the Bioethical Committee at the Universidad de Valparaíso.

Author Contributions: Soledad Torres and Laura Rifo developed the model and the methodology presented in this paper. Manuel Barahona performed the simulations and data analysis. Maritza Sepúlveda brought the applied problem and interpretations. All the authors have read and approved the final manuscript.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Harrison, J.M.; Shepp, L.A. On Skew Brownian motion. *Ann. Probab.* **1981**, *9*, 309–313.
2. Lejay, A. On the constructions of the Skew Brownian motion. *Probab. Surv.* **2006**, *3*, 413–466.
3. Kutoyants, Y. *Parameter Estimation for Stochastic Processes*; Heldermann: Berlin, Germany, 1984.
4. Lipster, R.; Shiryaev, A. *Statistics of Random Processes II, Applications*; Springer: Berlin/Heidelberg, Germany, 2001.
5. Florens-Zmirou, D. Statistics on crossings of discretized diffusions and local time. *Stoch. Process. Appl.* **1993**, *39*, 139–151.
6. Jacod, J. Rates of convergence to the local time of a diffusion. *Annales de l'IHP Probabilités et statistiques* **1998**, *34*, 505–544.
7. Bardou, O.; Martinez, M. Statistical estimation for reflected skew processes. *Stat. Inference Stoch. Process.* **2010**, *13*, 231–248.
8. Lejay, A.; Mordecki, E.; Torres, S. Is the Brownian motion skew? *Scand. J. Stat.* **2014**, *41*, 346–364.
9. Boness, D.J.; Bowen, W.D.; Oftedall, O.T. Evidence of a maternal foraging cycle resembling that of otariid seals in a small phocid, the harbor seal. *Behav. Ecol. Sociobiol.* **1994**, *34*, 95–104.
10. Karatzas, I.; Shreve, S. *Brownian Motion and Stochastic Calculus*; Springer: Berlin/Heidelberg, Germany, 1991.
11. Walsh, J.B. A diffusion with a discontinuous local time. *Astérisque* **1978**, *52*, 37–45.
12. Breiman, L. *Probability*; SIAM: Philadelphia, PA, USA, 1992.
13. Pereira, C.A.B.; Stern, J.M.; Wechsler, S. Can a significance test be genuinely bayesian? *Bayesian Anal.* **2008**, *3*, 15–36.
14. Rifo, L.; González-López, V. Full Bayesian analysis for a model of tail dependence. *Commun. Stat. Theory Methods* **2012**, *41*, 4107–4123.
15. Madruga, M.R.; Esteves, L.G.; Wechsler, S. On the bayesianity of Pereira–Stern tests. *Test* **2001**, *10*, 291–299.
16. Verdinelli, I.; Wasserman, L. Computing bayes factors using a generalization of the savage-dickey density ratio. *J. Am. Stat. Assoc.* **1995**, *90*, 614–618.
17. Bonadonna, F.; Lea, M.-A.; Dehorter, O.; Guinet, C. Foraging ground fidelity and route-choice tactics of a marine predator: The Antarctic fur seal *Arctocephalus gazella*. *Mar. Ecol. Prog. Ser.* **2001**, *223*, 277–286.
18. Robson, B.W.; Goebel, M.E.; Baker, J.D.; Ream, R.R.; Loughlin, T.R.; Francis, R.C.; Antonelis, G.A.; Costa, D.P. Separation of foraging habitat among breeding sites of a colonial marine predator, the northern fur seal (*Callorhinus ursinus*). *Can. J. Zool.* **2004**, *82*, 20–29.
19. Call, K.A.; Ream, R.R.; Johnson, D.; Sterling, J.T.; Towell, R.G. Foraging route tactics and site fidelity of adult female northern fur seal (*Callorhinus ursinus*) around the Pribilof Islands. *Deep Sea Res. II* **2008**, *55*, 1883–1896.
20. Riet-Sapriza, F.G.; Costa, D.P.; Franco-Trecu, V.; Marín, Y.; Chocca, J.; González, B.; Beathyate, G.; Chilvers, B.L.; Hückstadt, L.A. Foraging behavior of lactating South American sea lions (*Otaria flavescens*) and spatial-temporal resource overlap with the Uruguayan fisheries. *Deep Sea Res. II* **2013**, *88–89*, 106–119.

21. Rodríguez, D.H.; Dassis, M.; de León, A.P.; Barreiro, C.; Farenga, M.; Bastida, R.O.; Davis, R.W. Foraging strategies of Southern sea lion females in the La Plata River Estuary (Argentina-Uruguay). *Deep Sea Res. II* **2013**, *88–89*, 120–130.
22. Friedlaender, A.S.; Halpin, P.N.; Qian, S.S.; Lawson, G.L.; Wiebe, P.H.; Thiele, D.; Read, A.J. Whale distribution in relation to prey abundance and oceanographic processes in shelf waters of the Western Antarctic Peninsula. *Mar. Ecol. Prog. Ser.* **2006**, *317*, 297–310.
23. Leung, E.S.; Chilvers, B.L.; Nakagawa, S.; Moore, A.B.; Robertson, B.C. Sexual segregation in juvenile New Zealand sea lion foraging ranges: Implications for intraspecific competition, population dynamics and conservation. *PLoS ONE* **2012**, *7*, e45389.
24. Cappozzo, H.L.; Perrin, W.F. South American sea lion (*Otaria flavescens*). In *Encyclopedia of Marine Mammals*; Elsevier: Amsterdam, The Netherlands, 2008; pp. 1076–1079.
25. Sigler, M.F.; Tollit, D.J.; Vollenweider, J.J.; Thedinga, J.F.; Csepp, D.J.; Womble, J.N.; Wong, M.A.; Rehberg, M.J.; Trites, A.W. Steller sea lion foraging response to seasonal changes in prey availability. *Mar. Ecol. Prog. Ser.* **2009**, *388*, 243–261.