Parameter estimation from observations of first-passage times of the Ornstein-Uhlenbeck process and the Feller process

Susanne Ditlevsen
Department of Biostatistics, University of Copenhagen, DK 1014 Copenhagen K, Denmark
sudi@pubhealth.ku.dk

Ove Ditlevsen
Section of Maritime Engineering, Department of Mechanical Engineering,
Technical University of Denmark, DK 2800 Lyngby, Denmark
od@mek.dtu

Keywords: Biomedical stochastic processes, Chapman-Kolmogorov integral equation, Cox-Ingersoll-Ross process, diffusion processes, hidden random variation, parameter estimation from stopping times, stochastic neuronal models

ABSTRACT: Renewal point processes show up in many different fields of science and engineering. In some cases the renewal points become the only observable parts of an anticipated hidden random variation of some physical quantity. The hypothesis might be that a hidden random process originating from zero or some other low value only becomes visible at the time of first crossing of some given value level, and that the process is restarted from scratch immediately after the level crossing. It might then be of interest to reveal the defining properties of this hidden process from a sample of observed first-passage times. In this paper the hidden process is first anticipated as a non-stationary Ornstein-Uhlenbeck (OU) process with unknown parameters that have to be estimated only by use of the information contained in a sample of first-passage times. The estimation method is a direct application of the Fortet integral equation of the OU process. A non-stationary Feller process is considered subsequently. As the OU process the Feller process has a known transition probability distribution that allows the formulation of the integral equation. The described integral equation estimation method also provides a subjective graphical test of the applicability of the OU process or the Feller process when applied to a reasonably large sample of observed first-passage data. These non-stationary processes have several applications in biomedical research, for example as idealized models of the neuron membrane potential. When the potential reaches a certain threshold the neuron fires, whereupon the potential drops to a fixed initial value, from where it continuously builds up again until next firing. Also in civil engineering there are hidden random phenomena such as internal cracking or corrosion that after some random time break through to the material surface and become observable. However, the OU process has as a model of physical phenomena the defect of not being bounded to the negative side. This defect is not present for the Feller process, which therefore may provide a useful modeling alternative to the OU process.

1 INTRODUCTION

The first-passage time through a constant threshold of an Ornstein-Uhlenbeck process (OU) has been in focus for stochastic modeling of problems in which it seems as if a hidden random process only shows its existence when it reaches a certain level that triggers some observable event.

In physiology, neuronal firing activity has been modeled in this way, e.g. (Giorno, Lansky, Nobile, & Ricciardi 1988, Lansky, Sacerdote, & Tomasetti 1995, Wan & Tuckwell 1982). The potential difference that exists across the cell membrane, called the membrane potential, is modeled as an OU process acting between two consecutive neuronal firings. When the membrane voltage exceeds a certain threshold, the neuron releases a rapid electrical signal called a spike or an action potential. After firing, the membrane potential is reset to an initial voltage conveniently set to zero. Formally, the interspike interval is then identified with the first-passage time of the OU process through the threshold and the inter-spike intervals thus form a renewal process. The neurons communicate by sending these electrical impulses, and it is therefore of interest to understand the mechanisms behind the generation of spikes.

In survival analysis the OU process has been applied to describe a latent, unobserved development that influences the hazard of some "unit" under study (Aalen & Gjessing 2004). For instance, the process may represent the untreated development of a disease of a particular individual. When the process reaches a certain level, an event occurs for that individual, e.g. it may be diagnosis and start of treatment or it may be the death of the individual.

In mathematical finance the model has also been used, see e.g. (Alili, Patie, & Pedersen 2005, Linetsky 2004) and references therein.

In engineering there could be applications concerning hidden deterioration processes such as corrosion of reinforcement in concrete or internal cracking in materials. Since self-healing of material damages rarely is physically possible (but not excluded), the OU process $X(t)$ is physically unreasonable. However, it may with respect to first-passage properties still be applicable because the never decreasing physically possible process $\sup_{s\leq t} X(s)$ has the same first-passage properties as $X(t)$.

After treating the OU processes focus is on the Feller processes that have the physically healthy property of being bounded (a.s.) from below, a property that the OU processes as Gaussian processes do not possess. The one-dimensional distributions of those non-stationary Feller processes that become asymptotically stationary for large times are non-central χ^2-distributions with a time independent number of degrees of freedom and a time dependent non-centrality parameter that asymptotically approaches zero for large times (Feller 1951).

An extensive literature is dedicated to stochastic first-passage time problems. However, only few papers have appeared on how to estimate the parameters of a hidden random process from observations of its first-passage times even when the process is as simple as the OU process, e.g. (Ditlevsen & Lansky 2005, Inoue, Sato, & Ricciardi 1995). Herein we propose a method on how to estimate all free parameters, applicable both to the OU process and to the Feller process.

2 ORNSTEIN-UHLENBECK PROCESS

Consider the OU process X_t defined by the stochastic differential equation with initial condition

$$dX_t = \left(-\frac{X_t}{\tau} + \mu\right) dt + \sigma\, dW_t, \qquad X_0 = 0 \tag{1}$$

where W_t is a standard Wiener process, τ is a suitable time scale, and

$$T = \inf\{t > 0 : X_t \geq S\} \tag{2}$$

is the first-passage time of X_t through the constant threshold S. It is convenient to reformulate to the equivalent dimensionless form

$$d\left(\frac{X_t}{S}\right) = \left(-\frac{X_t}{S} + \frac{\mu\tau}{S}\right) d\left(\frac{t}{\tau}\right) + \frac{\sigma\sqrt{\tau}}{S}\, d\left(\frac{W_t}{\sqrt{\tau}}\right) \tag{3}$$

or

$$dY_s = (-Y_s + \alpha)\,ds + \beta\,dW_s, \qquad Y_0 = 0 \tag{4}$$

where

$$s = \frac{t}{\tau}, Y_s = \frac{X_t}{S}, W_s = \frac{W_t}{\sqrt{\tau}}, \alpha = \frac{\mu\tau}{S}, \beta = \frac{\sigma\sqrt{\tau}}{S} \tag{5}$$

and $T/\tau = \inf\{s > 0 : Y_s \geq 1\}$. Therefore, without loss of generality, all considerations in the following will be related to the dimensionless process Y_s and its first crossing of the level 1.

The solution process Y_s is normally distributed with conditional moments

$$E[Y_s|Y_0 = 0] = \alpha(1 - e^{-s}) \tag{6}$$

$$\text{Var}[Y_s|Y_0 = 0] = \frac{1}{2}\beta^2(1 - e^{-2s}) \tag{7}$$

The asymptotic mean is α and the asymptotic variance is $\beta^2/2$ as $s \to \infty$.

As suggested in (Wan & Tuckwell 1982) it is convenient to distinguish between two crudely defined parts of the parameter space called the suprathreshold regime and the subthreshold regime. The suprathreshold regime is defined as that for which $\alpha \gg 1$, that is, the asymptotic mean is far above the threshold implying that the first-passage times are relatively regular (deterministic behavior - which means that there are crossings also in the absence of noise). The subthreshold regime is defined as that for which $\alpha \ll 1$ implying that crossings of the threshold are caused only by random fluctuations (stochastic or Poissonian behavior). The term "Poissonian behavior" indicates that when $\alpha \ll 1$ (measured in terms of β), the first-passage times achieve characteristics of a Poisson point process (Nobile, Ricciardi, & Sacerdote 1985, Wan & Tuckwell 1982). What is left over of the parameter space is herein denoted as the threshold regime in which the asymptotic mean α does not deviate much from 1. The regimes are illustrated in Fig. 1.

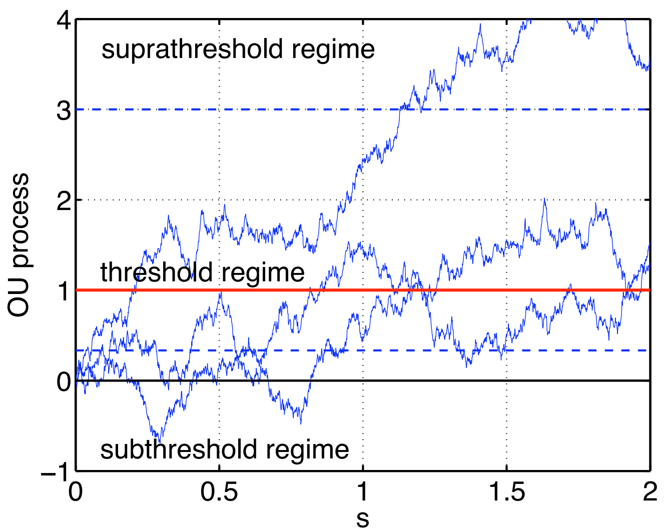

Figure 1. Illustration of the different regimes. The dashed line indicates the asymptotic mean α.

The goal of this paper is to present a method to estimate the unknown parameters α and β from independent observations of T/τ: $s_1, \ldots, s_N$, where N is the number of measurements. In neuronal modeling, the time scale

τ is an intrinsic parameter that characterizes the neuron irrespectively of the incoming signal (the time constant of the cell membrane), whereas μ and σ are input parameters describing signals that the neuron receives from surrounding neurons (Tuckwell & Richter 1978). The value of τ is often given by physiological considerations ($\approx 5-20$ msec), and it is only relevant as a guidance to assess a reasonable value range of the input parameters. As it is seen from the differential equation (3), the time scale τ can be freely chosen whereupon the parameters μ and σ are uniquely obtained from estimates of α and β. The estimates of α and β obtained by the method are finally compared with the results obtained from a previously proposed method for estimating the same parameters given that they are in the suprathreshold regime (Ditlevsen & Lansky 2005).

Let $f_{T/\tau}(s)$ be the probability density of T/τ. An exact expression for $f_{T/\tau}(s)$ is only known for $\alpha = 1$ and is

$$f_{T/\tau}(s)_{\alpha=1} = \frac{2e^{2s}}{\sqrt{\pi}\beta(e^{2s}-1)^{3/2}} \exp\Big(-\frac{1}{\beta^2(e^{2s}-1)}\Big) \tag{8}$$

Moreover, an asymptotically exact expression can be obtained as $\alpha \to \infty$ and $\beta \to \infty$ together under constant ratio α/β^2. Then the stochastic differential equation (4) asymptotically approaches the differential equation $\mathrm{d}Y_s = \alpha\,\mathrm{d}s + \beta\,\mathrm{d}W_s$ for a Wiener process with corresponding first-passage density

$$f_{T/\tau}(s)_{\alpha,\beta\to\infty} = \frac{1}{\sqrt{2\pi\beta^2 s^3}} \exp\Big(-\frac{(1-\alpha s)^2}{2\beta^2 s}\Big) \tag{9}$$

i.e., the inverse Gaussian distribution. If the limit passage is made by letting $\tau \to \infty$ with μ, σ, and S kept fixed, the corresponding asymptotic density function for T is

$$f_T(t)_{\tau\to\infty} = \frac{S}{\sqrt{2\pi\sigma^2 t^3}} \exp\Big(-\frac{(S-\mu t)^2}{2\sigma^2 t}\Big) \tag{10}$$

In these two special cases the maximum likelihood estimators of α and β^2 are

$$\alpha = 1: \qquad \check{\beta}^2 = \frac{1}{N}\sum_{i=1}^{N} \frac{2}{e^{2s_i}-1} \tag{11}$$

$$\alpha,\beta \to \infty: \quad \check{\alpha} = \frac{1}{\bar{s}} \quad \check{\beta}^2 = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{s_i} - \frac{1}{\bar{s}} \tag{12}$$

where $\bar{s} = \frac{1}{N}\sum_{i=1}^{N} s_i$. Otherwise no explicit expressions for the density are known, and other ways of estimating have to be found. One approach is to use moments or functions of moments of the first-passage time. In (Inoue, Sato, & Ricciardi 1995, Ricciardi & Sato 1988) series expressions for the moments of T were derived. In particular

$$E[T/\tau] = \frac{1}{2}\sum_{n=1}^{\infty}\frac{2^n}{n!}\frac{(1-\alpha)^n - (-\alpha)^n}{\beta^n}\Gamma\Big(\frac{n}{2}\Big) \tag{13}$$

However, this expression is difficult to work with, especially if $\alpha \gg 1$ (strongly suprathreshold) because of the canceling effects in the alternating series. In (Ditlevsen & Lansky 2005) the moments

$$E[e^{T/\tau}] = \frac{\alpha}{\alpha-1} \quad \text{if } \alpha > 1 \tag{14}$$

$$E[e^{2T/\tau}] = \frac{\alpha^2 - \beta^2/2}{(\alpha-1)^2 - \beta^2/2} \quad \text{if } \alpha - 1 > \frac{\beta}{\sqrt{2}} \tag{15}$$

are given. The first moment is only finite in the suprathreshold regime, and for the second moment to be finite it is further required that the asymptotic standard deviation of Y_s is smaller than the distance between the threshold and the asymptotic mean of Y_s.

Estimators for the parameters α and β^2 from observations of first-passage times in the suprathreshold regime are in (Ditlevsen & Lansky 2005) proposed as

$$\hat{\alpha} = \frac{Z_1}{Z_1 - 1}, \quad \hat{\beta}^2 = \frac{2(Z_2 - Z_1^2)}{(Z_2 - 1)(Z_1 - 1)^2} \tag{16}$$

where

$$Z_1 = \frac{1}{N}\sum_{i=1}^{N} e^{s_i}, \quad Z_2 = \frac{1}{N}\sum_{i=1}^{N} e^{2s_i} \tag{17}$$

based on the exact expressions (14) and (15). These simple estimators are only valid in suprathreshold regime, and it remains to determine the regime from data, a problem which is not straightforward.

3 FORTET INTEGRAL EQUATION

The stationary OU process solution to the stochastic differential equation (4) has the mean value α, variance $\beta^2/2$, and correlation function $\exp(-s)$. By introducing the condition that $Y_0 = 0$, a non-stationary Gaussian process with mean value function $E[Y_s \mid Y_0 = 0] = [1 - \exp(-s)]\alpha$, and variance function $\text{Var}[Y_s \mid Y_0 = 0] = [1 - \exp(-2s)]\beta^2/2$ is obtained. In the following we will denote this non-stationary process by $Y(s)$ without explicitly mentioning the condition $Y_0 = 0$.

Let $f(s)$ be the density function for the time t/τ from zero to the first crossing of the level 1 by Y. The probability

$$P[Y_s > 1] = \Phi\Big(\frac{\alpha(1 - e^{-s}) - 1}{\sqrt{1 - e^{-2s}}\,\beta/\sqrt{2}}\Big) \tag{18}$$

where $\Phi(\cdot)$ is the standard normal distribution function, can alternatively be calculated by the transition integral

$$P[Y_s > 1] = \int_0^s f(u)\,\Phi\bigg(\frac{\alpha-1}{\beta/\sqrt{2}}\frac{1-e^{-(s-u)}}{\sqrt{1-e^{-2(s-u)}}}\bigg)\,\mathrm{d}u \tag{19}$$

because

$$P[Y_s > 1 \mid Y_u{=}1] = \Phi\Big(\frac{E[Y_s \mid Y_u{=}1] - 1}{D[Y_s \mid Y_u]}\Big)$$

$$= \Phi\bigg(\frac{\alpha + e^{-(s-u)}(1 - \alpha) - 1}{\beta/\sqrt{2}\sqrt{1-e^{-2(s-u)}}}\bigg) \tag{20}$$

where $D[\cdot|\cdot]$ indicates conditional standard deviation. Thus we have obtained an integral equation for the unknown density function $f(s)$. It is difficult to obtain a solution except for the case $\alpha = 1$, where the factor to $f(u)$ under the integral sign becomes the constant 1/2. Then the distribution function $F(s)$ of the first-passage time simply becomes

$$F(s) = 2\Phi\Big(\frac{-e^{-s}\sqrt{2}}{\beta\sqrt{1 - e^{-2s}}}\Big), \quad s > 0 \tag{21}$$

and by differentiation

$$f(s) = \frac{e^{-s}\,2\sqrt{2}}{\beta(1-e^{-2s})^{3/2}}\varphi\Big(\frac{e^{-s}\sqrt{2}}{\beta\sqrt{1-e^{-2s}}}\Big) \tag{22}$$

where $\varphi(\cdot)$ is the standard normal density function. This is the density (8).

If α and β are made large under constant ratio α/β^2 it is seen that the range of s in which there is essential variation of the right side of (18) narrows down closer and closer to zero implying that $\alpha(1-e^{-s}) \approx \alpha s$ and $\sqrt{1-e^{-2s}}\,\beta/\sqrt{2} \approx \beta\sqrt{s}$ asymptotically as α and β increase. Thus the integral equation takes the limit form

$$\Phi\Big(\frac{\alpha s-1}{\beta\sqrt{s}}\Big) = \int_0^s f(u)\,\Phi\left(\frac{\alpha\sqrt{s-u}}{\beta}\right)\,\mathrm{d}u \tag{23}$$

which by differentiation with respect to s gives the equation

$$\frac{\alpha s+1}{2\beta s\sqrt{s}}\varphi\Big(\frac{\alpha s-1}{\beta\sqrt{s}}\Big)$$

$$= \frac{1}{2}f(s) + \frac{\alpha}{\beta}\int_0^s \frac{f(u)}{2\sqrt{s-u}}\varphi\Big(\frac{\alpha\sqrt{s-u}}{\beta}\Big)\,\mathrm{d}u \tag{24}$$

It is not an easy task to solve this integral equation directly. However, we may check whether the density function (9) satisfies the equation. Indeed, it is the case. The calculations are shown in Appendix 1.

4 PARAMETER ESTIMATION

The integral equation

$$\Phi\Big(\frac{\alpha(1-e^{-s})-1}{\sqrt{1-e^{-2s}}\,\beta/\sqrt{2}}\Big)$$

$$= \int_0^s f(u)\,\Phi\left(\frac{\alpha-1}{\beta/\sqrt{2}}\sqrt{\frac{1-e^{-(s-u)}}{1+e^{-(s-u)}}}\right)\mathrm{d}u \tag{25}$$

obtained from (18) and (19) can be used to estimate the two parameters α and β as explained in the following. The resulting estimators do not depend on the regime.

Consider the ordered sample $s_1 \le s_2 \le \ldots \le s_N$ of independent observations of T/τ. For given values of the parameters the right hand side of (25) can be estimated at s from the sample by the average

$$\mathrm{RHS}(s) \approx \frac{1}{N}\sum_{i=1}^{\max\{n:s_n\le s\}} \Phi\left(\frac{\alpha-1}{\beta/\sqrt{2}}\sqrt{\frac{1-e^{-(s-s_i)}}{1+e^{-(s-s_i)}}}\right) \tag{26}$$

since it is the expected value of

$$1_{U\in[0,s]}\Phi\left(\frac{\alpha-1}{\beta/\sqrt{2}}\sqrt{\frac{1-e^{-(s-U)}}{1+e^{-(s-U)}}}\right) \tag{27}$$

with respect to the distribution of $U = T/\tau$, where $1_{U\in[0,s]}$ is the indicator function for the event $\{U \in [0,s]\}$.

Upon dividing both sides of the integral equation (25) by $\Phi[(\alpha-1)/(\beta/\sqrt{2})]$, the two sides become identical distribution functions in s when $\alpha \ge 0$. In this case the right hand side of (26) gets (except for normalization) the form as a statistical fluctuating empirical distribution function corresponding to the theoretical distribution function on the left hand side of (25). As in the well-known Kolmogorov-Smirnov test a convenient measure of the statistical error is therefore the max norm

$$L(\alpha, \beta) = \max_{s \in \mathbb{R}_+} |\mathrm{RHS}_{\mathrm{emp}}(s) - \mathrm{LHS}(s)| \tag{28}$$

where $\mathrm{RHS}_{\mathrm{emp}}(s)$ is the right hand side of (26) and $\mathrm{LHS}(s)$ is the left hand side of (25). The estimates of α and β are then directly obtained by minimizing this error function, i.e.,

$$(\tilde{\alpha}, \tilde{\beta}) = \arg\min_{\alpha,\beta} L(\alpha, \beta) \tag{29}$$

Numerically the search for the maximum in (28) must be restricted to a finite set of s-values, of course. A good choice of this set is obtained as $\{s \in \mathbb{R}_+ : LHS(s)/\max LHS = i/N, i = 1, \ldots, N-1\}$ for some reasonably large number N ($N = 100$, say).

As given by the left side of the integral equation, the probability of being above the threshold at time s takes the maximal value $\Phi[-\sqrt{1-2\alpha}/(\beta/\sqrt{2})]$ for $s = \log(1-1/\alpha)$ if $\alpha < 0$. This non-monotonic behavior has no influence on the parameter estimation by minimization of the max-norm of the deviation between the two sides of the integral equation.

A possible set of initial parameter values for the optimization procedure can be obtained from (16) for α, and (12) for β.

5 NUMERICAL RESULTS

Trajectories from the OU process were simulated according to the Euler scheme for the same $\beta = 1$ in all cases but for different $\alpha = 0.8$ (subtreshold), 1 (threshold), 2, 3, 4 (suprathreshold) and 11 (strongly suprathreshold, approximating the Wiener process). The time step size was 0.001, except for $\alpha = 11$ where the step size were 0.0001 because the crossings in this case are more dense in time. For each pair of parameter values, the process was run until the trajectory reached the threshold value 1 where the time was recorded. This was repeated 10.000 times, and the recorded times were separated in 100 samples with 100 observations in each sample. These 100 samples were used to obtain 100 estimates of the parameter pair (α, β).

To compare the different methods, (α, β) was estimated by the maximum likelihood estimator (11) in the threshold regime, by the maximum likelihood estimator (12) for the Wiener approximation, by the moment method (16) in the threshold and the suprathreshold regime, and finally by the estimator (29) derived from the integral equation. Only the last estimator is valid in the entire parameter space. The problem of determining the regime from the data is not treated. It is therefore assumed that the regime is known before applying the first estimators. In most applications such knowledge is not available.

The estimation results for α are listed in Table 1. The estimator (29) works well in all cases with the important advantage that it does not depend on the regime. In the suprathreshold regime the α-estimator (16) performs slightly better, but requires knowledge of the regime. On the other hand, the α-estimator (16) is easier to implement and is thus preferable if the regime is known to be suprathreshold. The estimation results for β are

Table 1. Values of α and β used in the simulations of OU realizations, and the corresponding averages of the samples of 100 parameter estimates of α and β $\pm$ the sample standard deviation (SSD) using (16) and (29). All first-passage time samples contain 100 simulated observations each. Besides the estimates in the table the estimates obtained from (11) and (12) are $\breve{\beta} = 0.97 \pm 0.0.07$ for $\alpha = 1$ and $\breve{\alpha} = 10.32 \pm 0.36$, $\breve{\beta} = 0.98 \pm 0.0.07$ for the approximate Wiener process case.

regime	$\beta = 1$ $\alpha =$	statistics of 100 estimates: average $\pm$ SSD $\hat{\alpha}$	$\tilde{\alpha}$
subthr.	0.8		0.83$\pm$ 0.14
threshold	1	1.14 $\pm$ 0.08	1.01 $\pm$ 0.15
suprathr.	2	1.99 $\pm$ 0.13	1.97 $\pm$ 0.15
suprathr.	3	3.00 $\pm$ 0.17	2.99 $\pm$ 0.17
suprathr.	4	3.96 $\pm$ 0.20	3.94 $\pm$ 0.21
$\approx$ Wiener	11	10.78 $\pm$ 0.36	10.76 $\pm$ 0.37
		$\hat{\beta}$	$\tilde{\beta}$
subthr.	0.8		0.97$\pm$ 0.10
threshold	1		0.98$\pm$ 0.11
suprathr.	2	0.78 $\pm$ 0.09	0.98 $\pm$ 0.09
suprathr.	3	0.95 $\pm$ 0.13	0.98 $\pm$ 0.09
suprathr.	4	0.95 $\pm$ 0.11	1.00 $\pm$ 0.09
$\approx$ Wiener	11	0.98 $\pm$ 0.09	0.99 $\pm$ 0.09

also listed in Table 1. The estimator (29) performs well in all cases. Only in the strong suprathreshold regime the β-estimator (16) performs as well. In all cases the estimator (29) is preferable.

Plots of the marginal sample distribution functions of the 100 estimate pairs obtained by the estimator (29) indicate that the estimator can be assumed to be normally distributed marginally, and it is thus straightforward to construct confidence intervals. This assumption is quite good as seen from Fig. 2. Moreover, the estimator is practically unbiased, maybe except for the strongly suprathreshold regime, where it shows a slight bias. This is seen in the diagrams of Figs. 8 and 9 with the scatterplots of the estimated pairs (α, β) for all simulated cases. Assuming that the marginal distributions are normal it is by inspection of the diagrams not unreasonable to adopt a binormal distribution model for the estimator pair $(\tilde{\alpha}, \tilde{\beta})$. It is noted that the two estimators are correlated except in the high suprathreshold regime.

That the joint distribution of the two estimators can be taken as binormal is anticipated as a large sample asymptotic result sufficiently accurate for sample sizes at least as large as investigated here (100 first-passage times). How small the sample size can be before the normal distribution assumption becomes too inaccurate remains to be investigated. In any specific application on real measured data a simulation study should be made to allow statements about the estimation accuracy. Herein the small sample behavior of the estimators is investigated by subdividing the data set of the 10 000 simulated first-passage times for $\alpha = 2$ and $\beta = 1$. The large sample is separated into 1000 samples with only 10 observations in each sample and (α, β) is estimated for each of these. For four of these the procedure did not converge. The 996 estimate pairs are plotted in Fig. 3. The scatter plot indicates that even for a sample size as small as 10 it seems reasonable to adopt the binormal distribution for the estimator pair.

The integral equation estimation method was also applied to data sets generated with $\beta = 2$ for the same values of α as for $\beta = 1$ (results not shown). For this larger variance of the underlying process, the method tends to underestimate α slightly in the threshold and the suprathreshold regime. The estimates for β are in all

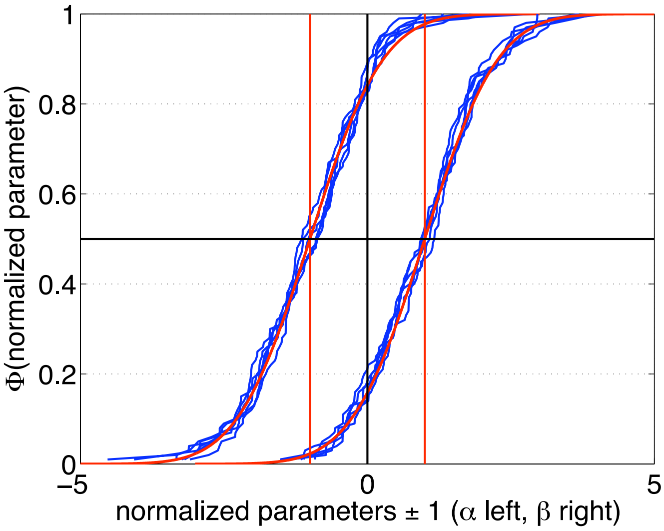

Figure 2. Normalized empirical distribution functions of the sample of 100 joint estimates of α and β compared to the standardized normal distribution function.

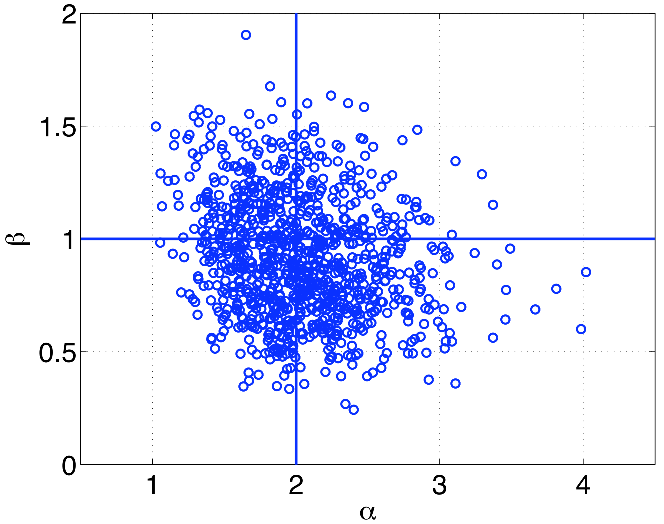

Figure 3. Scatterplots of the 996 pairs of estimates of (α, β), each estimated from a sample of 10 simulated first-passage times corresponding to the true values $\alpha = 2$ and $\beta = 1$.

cases downward biased. As expected, the sample standard deviations of the estimates are larger than for β=1 (approximately twice as large for the estimates of α, and one and a half times as large for the estimates of β).

6 FELLER PROCESS

In many applications the OU process is physically unrealistic particularly because it is unbounded. Even though this can be artificially mended by using the process $\sup_{u \leq s} Y(u)$, a more reasonable model candidate is the Feller process (Feller 1951), also called the Cox-Ingersoll-Ross process in the financial literature (Cox, Ingersoll, & Ross 1985). The Feller process is bounded from below, a property which is introduced in neuronal modeling by the action of an inhibitory reversal potential. The process is in its dimensionless form defined by the stochastic differential equation

$$\mathrm{d}Y_s = (-Y_s + \alpha)\ \mathrm{d}s + \frac{\beta}{\sqrt{\alpha}}\sqrt{Y_s}\,\mathrm{d}W_s \tag{30}$$

In case $2(\alpha/\beta)^2 \geq 1$ the process stays positive at all times (a.s.) (in Feller's terminology the boundary at zero is a so-called entrance boundary defined as a boundary from which the process can start but not return to (Karlin & Taylor 1981), p.234, and the Feller process becomes asymptotically stationary as $s \to \infty$. An illustration of a sample path is shown in Fig. 4. Feller has shown that the transition probability distribution is a non-central χ^2-distribution with $\nu = 4(\alpha/\beta)^2$ degrees of freedom and with conditional mean

$$E[Y_s|Y_0 = y_0] = \alpha + (y_0 - \alpha)e^{-s} \tag{31}$$

and variance

$$\mathrm{Var}[Y_s|Y_0 = y_0] = \frac{\beta^2}{2}(1 - e^{-s})\Big[1 + \Big(\frac{2y_0}{\alpha} - 1\Big)e^{-s}\Big] \tag{32}$$

see e.g. (Cox, Ingersoll, & Ross 1985). The standard form of the non-central χ^2-distribution with ν degrees of freedom and non-centrality parameter δ has mean $\nu + \delta$ and variance $2\nu + 4\delta$. The conditional mean (31) and conditional variance (32) are both obtained if the affinity factor $1/a(s) = \beta^2(1 - e^{-s})/(4\alpha)$ is applied to the standardized variable and the non-centrality parameter is $\delta(s, y_0) = (4\alpha y_0/\beta^2)[e^{-s}/(1 - e^{-s})]$. Thus

$$P(Y_s \leq y \mid Y_0 = y_0) = F_{\chi^2}[a(s)y, \nu, \delta(s, y_0)] \tag{33}$$

where $F_{\chi^2}(x, \nu, \delta)$ is the standard non-central χ^2-distribution function of ν degrees of freedom and non-centrality parameter δ [in *Matlab* denoted by ncx2cdf(x,ν,δ)]. It is seen that $\delta \to 0$ as $s \to \infty$, i.e., the asymptotic distribution is a gamma distribution.

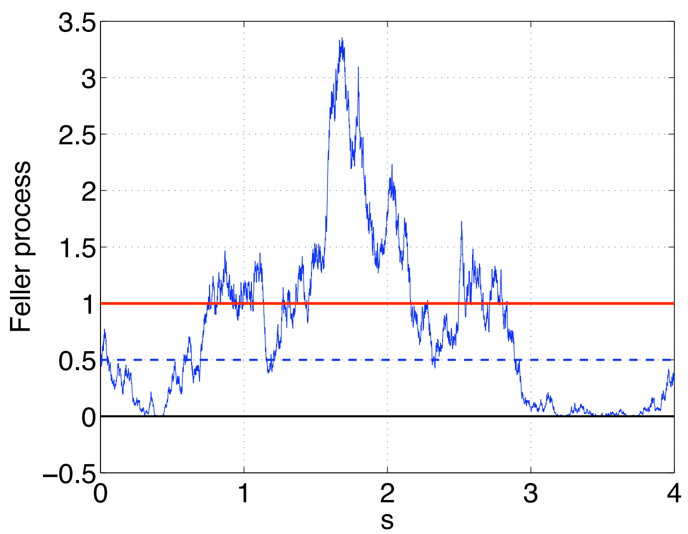

Figure 4. Example of realization of Feller process.

With these results the Fortet integral equation for the probability density $f(s)$ of the first-passage time T/τ through the level 1 by the non-stationary Feller process $Y_s \mid Y_0 = y_0 < 1$ is directly obtained as

$$1 - F_{\chi^2}[a(s), \nu, \delta(s, y_0)] \tag{34}$$

$$= \int_0^s f(u)\{1 - F_{\chi^2}[a(s-u), \nu, \delta(s-u, 1)]\}\,\mathrm{d}u$$

As for the OU process, the parameters α, β, and here also the initial value y_0 (if not known in advance) can be estimated from a suitably large sample of first passage times of the Feller process by interpreting the right hand side of (34) as an expectation. The solid lines in the diagrams of Fig. 5 show the good fit between the estimated two sides of (34) with $y_0 = 0.5$ known. As starting values in the optimization procedure (39) can be used for α and (12) for β^2/α.

Series expressions for the mean of T are given in (Giorno, Lansky, Nobile, & Ricciardi 1988, Lansky, Sacerdote, & Tomasetti 1995):

$$E[T/\tau] = \sum_{n=0}^{\infty} \frac{1 - y_0^{n+1}}{(n+1)\prod_{i=0}^{n}(\alpha + i\beta^2/2\alpha)} \tag{35}$$

(Ditlevsen & Lansky 2006) give the moments

$$E[e^{T/\tau}] = \frac{\alpha - y_0}{\alpha - 1} \quad \text{if} \quad \alpha > 1 \tag{36}$$

$$E[e^{2T/\tau}] = \frac{2\alpha(\alpha - y_0)^2 + \beta^2(\alpha - 2y_0)}{2\alpha(\alpha - 1)^2 + \beta^2(\alpha - 2)} \quad \text{if} \tag{37}$$

$$\sqrt{1 + 2(\alpha/\beta)^2} < 1 + 2\alpha(\alpha - 1)/\beta^2 \tag{38}$$

that otherwise are infinite. Assuming that the data are in the allowed parameter region, moment estimators of the parameters are then obtained from equations (36) and (37) as

$$\hat{\alpha} = \frac{Z_1 - y_0}{Z_1 - 1} \tag{39}$$

and

$$\hat{\beta}^2 = \frac{2(1 - y_0)^2(Z_2 - Z_1^2)}{2(Z_1-1)(Z_2-y_0) - (Z_1-y_0)(Z_2-1)}\hat{\alpha} \tag{40}$$

where Z_1 and Z_2 are given by (17).

Just as for the OU process 100 times 100 simulations are run for the Feller process. The statistics of the estimates of α and β are given in Table 2. The quality and distribution properties are about the same as for the OU process.

The less good accuracy corresponding to $\alpha = 11$ are due to numerical difficulties in connection with computation of χ^2-distribution values.

7 MODEL TESTING

With the parameter estimates from samples of 100 first-passage times of the OU process simulated for the selected values of α and for β=1, the left side of the integral equation (25) is in Fig. 5 compared to the right side as estimated by the respective samples.

The diagrams in Fig. 6 show the error of fit in terms of the difference between the two normalized sides of the integral equation (34) (solid line) and of the integral equation (25) (dashed line) corresponding to the

Table 2. Values of α and β used in the simulations of Feller process realizations, and the corresponding averages of the samples of 100 parameter estimates of α and β $\pm$ the sample standard deviation (SSD), using (39), (40) and (29), where relevant. All first-passage time samples contain 100 simulated observations each.

regime	$\beta = 1$ $\alpha =$	statistics of 100 estimates: average $\pm$ SSD $\hat{\alpha}$	$\tilde{\alpha}$
subthr.	0.8		0.79$\pm$ 0.09
threshold	1	1.10 $\pm$ 0.06	1.00$\pm$ 0.08
suprathr.	2	1.99 $\pm$ 0.10	1.98 $\pm$ 0.11
suprathr.	3	2.97 $\pm$ 0.09	2.95 $\pm$ 0.10
suprathr.	4	3.94 $\pm$ 0.12	3.90 $\pm$ 0.11
$\approx$ Wiener	11	10.96 $\pm$ 0.11	9.88 $\pm$ 0.15
		$\hat{\beta}$	$\tilde{\beta}$
subthr.	0.8		0.94$\pm$ 0.10
threshold	1		0.93$\pm$ 0.10
suprathr.	2	0.64 $\pm$ 0.10	0.95 $\pm$ 0.08
suprathr.	3	0.48 $\pm$ 0.05	0.96 $\pm$ 0.10
suprathr.	4	0.39 $\pm$ 0.04	0.92 $\pm$ 0.09
$\approx$ Wiener	11	0.22 $\pm$ 0.02	1.46 $\pm$ 0.16

parameter value estimates of α and β obtained for the two processes on the basis of the same samples of 100 first-passage times simulated from the Feller process. The error size and variation with time is very closely the same for the two integral equations even though the parameter estimates, and thus the asymptotic means and standard deviations, come out rather different in value. The conclusion that can be drawn from the diagrams in Fig. 6 is that a graphical test of the validity of the OU model as compared to the Feller model and vice versa is difficult to construct only on the basis of the sample of first-passage times from the one or the other process. Obviously this is because the family of first-passage time distributions of both processes are rather flexible with respect to fitting to data. Consequently physical reasons or other types of data than first-passage time samples must be found for supporting the choice.

There are renewal processes that deviate sufficiently from those generated by threshold passages of the OU process or the Feller process to reveal by the graphical test that these processes are not plausible as underlying processes. This is illustrated in the diagrams of Fig. 7.

CONCLUSIONS

The investigation herein is about estimation of the parameters of a random process defined by a given stochastic differential equation when the process is hidden except for observation of times to first passages from below of a given threshold. In particular the investigation focuses on a non-stationary Ornstein-Uhlenbeck process and a non-stationary Feller process for which the Fortet integral equations for the first-passage time probability density are explicitly known. It is shown by simulations that the two parameters of the processes can be quite accurately estimated by applying the integral equations to suitably large samples of first-passage times, and empirically that the bivariate normal distribution is applicable for the correlated pair of estimates. Some other available estimation procedures are applied to the simulated samples with the conclusion that the integral equation method

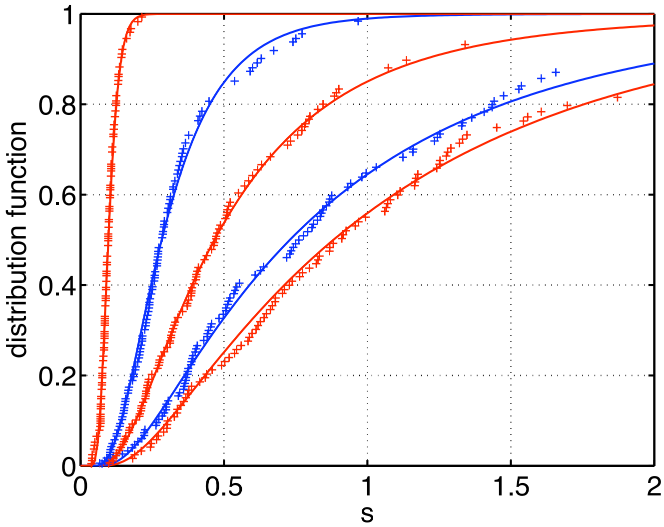

Figure 5. Comparison of the (normalized) left hand side of the integral equation (25) (smooth curves) to the empirical (normalized) right hand side given by (26) for five simulated samples of 100 first-passage times of the OU process of the level 1 corresponding to the true α-values 1, 2, 3, 4, 11, respectively, and the true $\beta = 1$ (right to left). For these samples the estimates of (α, β) according to (29) are (1.212, 0.926), (1.677, 0.996), (2.657, 1.039), (4.055, 1.029), (10.801, 0.956), respectively.

generally gives at least as good or more accurate estimates. Moreover the integral equation method is more generally applicable than the other available estimation methods.

When it comes to testing of whether the one or the other process type best models the hidden process it is demonstrated that a sample of first-pasage times can originate from any of the two process models without seeing any significant difference in the goodness of fit. Thus other reasons must be used to guide the choice of model. However, for many other renewal processes it can be seen that they are not well compatible with any of the two processes.

Acknowledgments This work is supported by grants from the Danish Medical Research Council and the Lundbeck Foundation to S. Ditlevsen.

APPENDIX 1

By substitution of the density function (9) in the integral in (24) we get

$$\frac{\alpha}{\beta}\int_0^s \frac{f(u)}{2\sqrt{s-u}}\varphi\left(\frac{\alpha\sqrt{s-u}}{\beta}\right)\mathrm{d}u = \frac{\alpha}{\beta}\frac{1}{4\pi\beta}\times$$

$$\int_0^s \frac{1}{u\sqrt{u(s-u)}}\exp\left(-\frac{(1-\alpha u)^2+\alpha^2(s-u)u}{2\beta^2 u}\right)\mathrm{d}u$$

$$= \frac{\alpha}{4\pi\beta^2}\exp\left(\frac{(2-\alpha s)\alpha s}{2\beta^2 s}\right)\int_0^s \frac{\exp\left(-1/(2\beta^2 u)\right)}{u\sqrt{u(s-u)}}\mathrm{d}u$$

$$= \frac{\alpha}{4\pi\beta^2 s}\exp\left(\frac{(2-\alpha s)\alpha s}{2\beta^2 s}\right)\int_0^1 \frac{e^{-\gamma/w}}{w\sqrt{w(1-w)}}\mathrm{d}w$$

where $\gamma = 1/(2\beta^2 s)$ and where the substitution $w = u/s$ is used. The definite integral is manipulated further as follows:

$$\int_0^1 \frac{e^{-\gamma/w}}{w\sqrt{w(1-w)}}\,\mathrm{d}w = \int_1^\infty \frac{e^{-\gamma v}}{\sqrt{v-1}}\,\mathrm{d}v =$$

$$2e^{-\gamma}\int_0^\infty e^{-\gamma u}\,\mathrm{d}\sqrt{u} = \frac{2e^{-\gamma}}{\sqrt{\gamma}}\int_0^\infty e^{-\frac{1}{2}(\sqrt{2\gamma u})^2}\,\mathrm{d}\sqrt{\gamma u}$$

$$= \sqrt{\frac{\pi}{\gamma}}e^{-\gamma} = \sqrt{2\pi}\beta\sqrt{s}\exp\Big(-\frac{1}{2\beta^2 s}\Big)$$

where the substitutions $v = 1/w$ and $u = v - 1$ are applied. Thus

$$\frac{\alpha}{\beta}\int_0^s \frac{f(u)}{2\sqrt{s-u}}\varphi\Big(\frac{\alpha\sqrt{s-u}}{\beta}\Big)\,\mathrm{d}u = \frac{\alpha s}{2\beta s\sqrt{s}}\,\varphi\Big(\frac{\alpha s-1}{\beta\sqrt{s}}\Big)$$

Adding $f(s)/2$ we obtain the left side of the integral equation (24). This concludes the proof.

REFERENCES

Aalen, O. & H. Gjessing (2004). Survival models based on the Ornstein-Uhlenbeck process. *Lifetime data analysis* 10:407–423.

Alili, L., P. Patie, & J. Pedersen (2005). Representations of the first hitting time density of an Ornstein-Uhlenbeck process. *Stochastic Models* 21:967–980.

Cox, J., J. Ingersoll, & S. Ross (1985). A theory of the term structure of interest rates. *Econometrica* 53:385–407.

Ditlevsen, S. & P. Lansky (2005). Estimation of the input parameters in the Ornstein-Uhlenbeck neuronal model. *Phys. Rev. E* 71(011907).

Ditlevsen, S. & P. Lansky (2006). Estimation of the input parameters in the Feller neuronal model. *Phys. Rev. E* 73(061910).

Feller, W. (1951). Diffusion processes in genetics. In J. Neyman (ed.), *Proceedings of the second Berkeley symposium on mathematical statistics and probability*, pp. 227–246. University of California Press, Berkeley.

Giorno, V., P. Lansky, A. Nobile, & L. Ricciardi (1988). Diffusion approximation and first-passage-time problem for a model neuron. *Biol. Cybern.* 58:387–404.

Inoue, J., S. Sato, & L. Ricciardi (1995). On the parameter estimation for diffusion models of single neurons activity. *Biol. Cybern.* 73:209–221.

Karlin, S. & H. Taylor (1981). *A second course in stochastic processes*. Academic Press, San Diego.

Lansky, P., L. Sacerdote, & F. Tomasetti (1995). On the comparison of Feller and Ornstein-Uhlenbeck models for neural activity. *Biol. Cybern.* 73:457–465.

Linetsky, V. (2004). Computing hitting time densities for CIR and OU diffusions: applications to mean-reverting models. *J. Comput. Finance* 7:1–22.

Nobile, A., L. Ricciardi, & L. Sacerdote (1985). Exponential trends of Ornstein-Uhlenbeck 1st-passage-time densities. *J. Appl. Prob.* 22:360–369.

Ricciardi, L. & S. Sato (1988). First-passage-time density and moments of the Ornstein-Uhlenbeck process. *J. Appl. Prob.* 25:43–57.

Tuckwell, H. & W. Richter (1978). Neuronal interspike time distributions and the estimation of neurophysiological and neuroanatomical parameters. *J. Theor. Biol.* 71:167–180.

Wan, F. & H. Tuckwell (1982). Neuronal firing and input variability. *J. Theoret. Neurobiol.* 1:197–218.

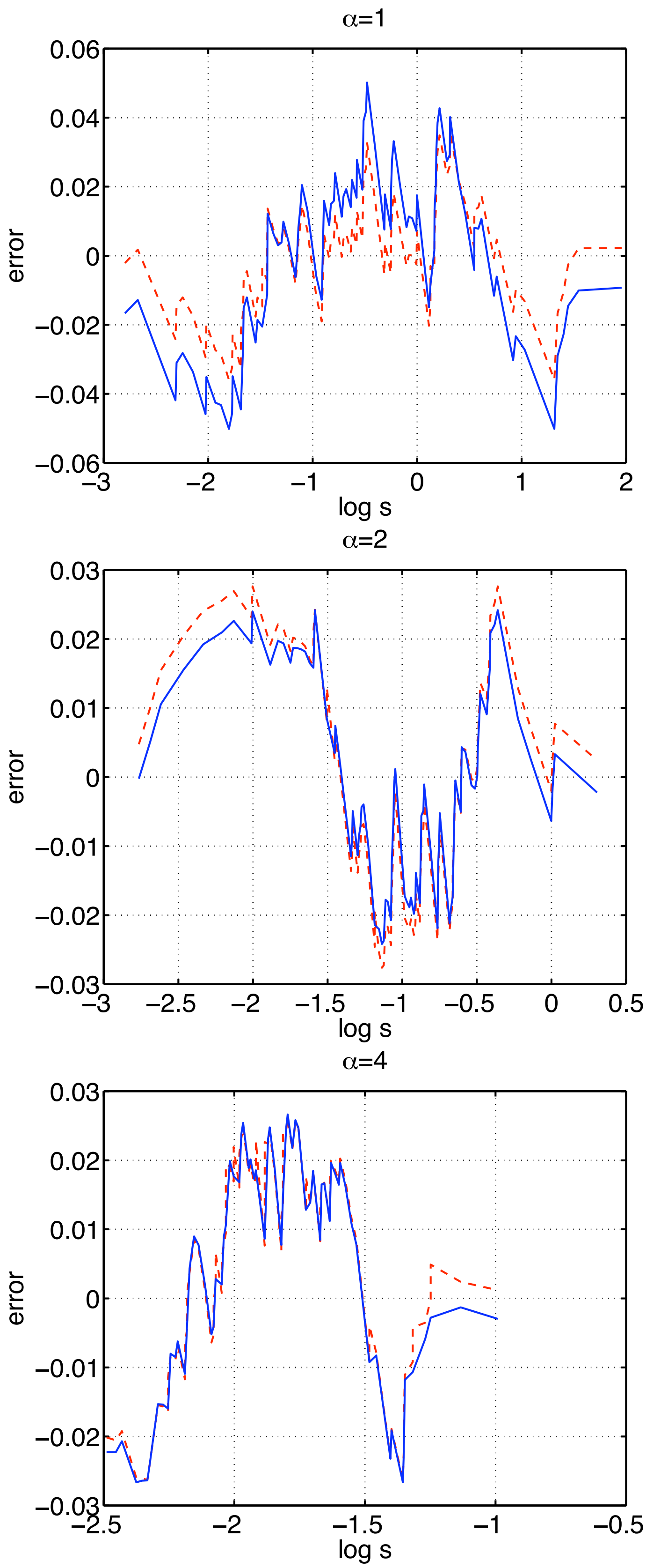

Figure 6. The fitting error LHS-RHS (LHS, RHS are normalized to the range [0, 1]) when estimating the parameters of the Feller process by use of (34) (solid line) or the parameters of the OU process by use of (25) (dashed line) from a simulated sample of 100 first-passage times of the level 1 by the Feller process with the α-value written on top of each diagram and all for $\beta = 1$ and $y_0 = 0.5$. The errors are close to be the same for the two models.

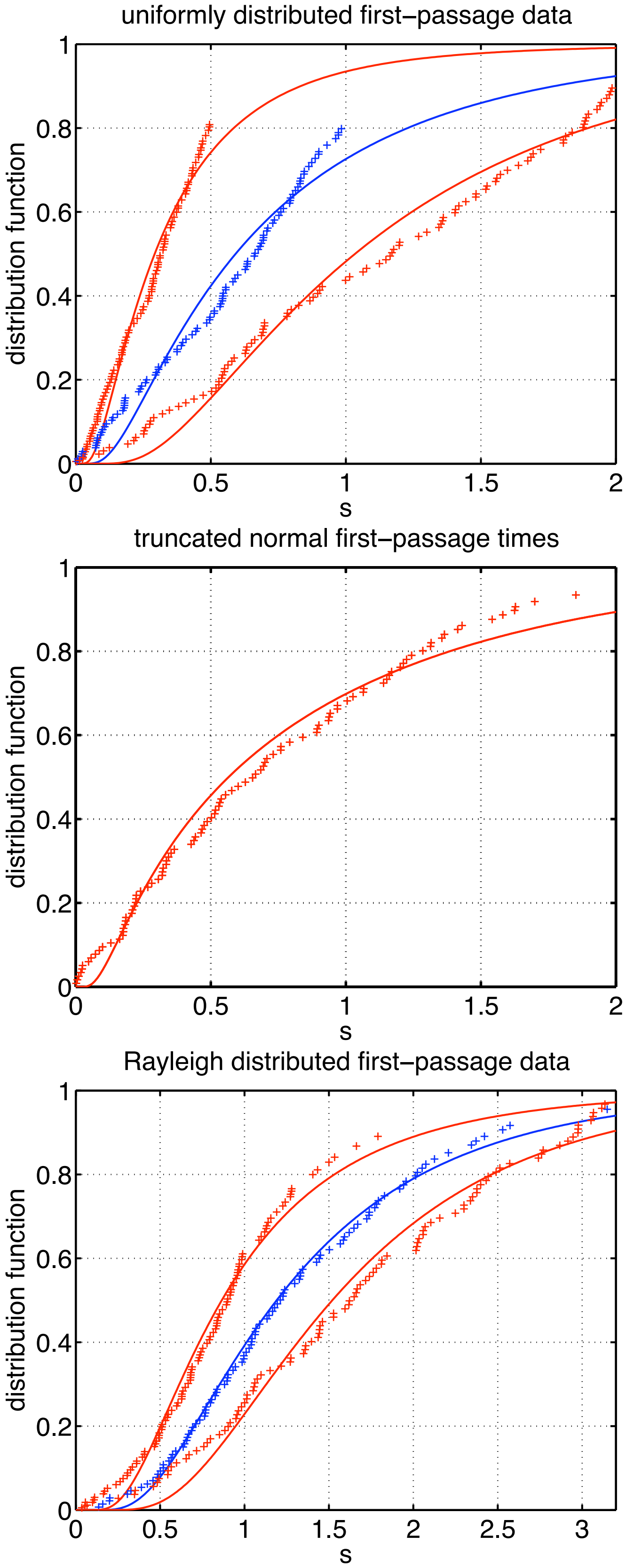

Figure 7. Indication of misfit of the OU model when the first-passage data are uniformly distributed (top),truncated normally distributed (density $\propto \varphi(s)$, $s > 0$) (middle), and Rayleigh distributed (density $\propto s\varphi(s/\sigma)$, $s > 0$, $\sigma^2 = 0.5.1, 2$ from left to right). The test uses the integral equation (25).

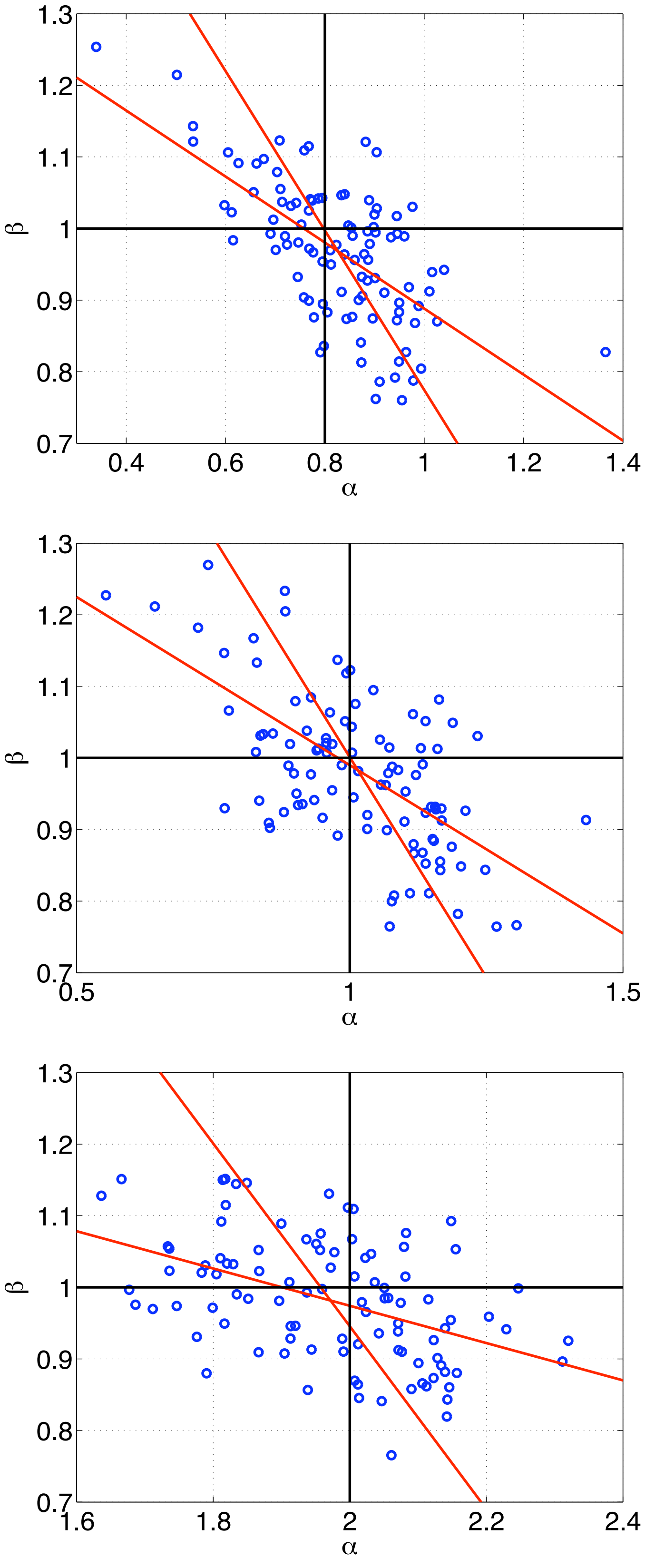

Figure 8. This figure continues in next figure, see caption in Fig. 9.

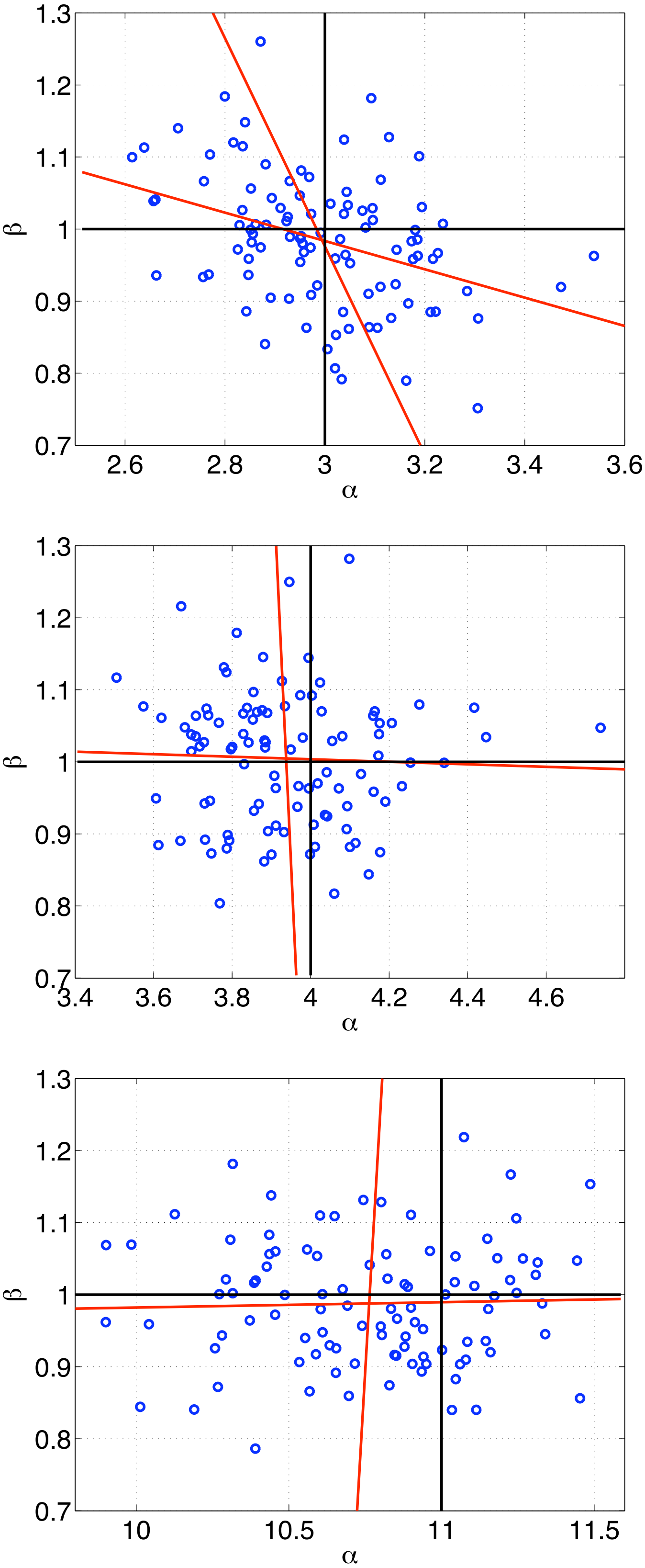

Figure 9. Scatterplots of the 100 pairs of estimates of (α, β), each estimated from a sample of 100 simulated first-passage times for the OU process. From the top the simulated data correspond to the true values $\alpha = 0.8, 1, 2, 3, 4, 11$ and $\beta = 1$, respectively. The shown axis parallel lines cross at the point of true parameter values. The two other lines are the estimated linear regressions of β on α and α on β, illustrating the degree of correlation between the two estimators. The estimated correlation coefficients are -0.643, -0.618, -0.452, -0.369, -0.039, 0.032, respectively.