IET Radar, Sonar & Navigation

Research Article

IET Journals
The Institution of Engineering and Technology

Estimation of the parameters of stochastic differential equations for sea clutter

ISSN 1751-8784
Received on 25th September 2018
Accepted on 18th October 2018
doi: 10.1049/iet-rsn.2018.5445
www.ietdl.org

Clement Julien Roussel[1] ✉, *Arnaud Coatanhay*[1], *Alexandre Baussard*[1]

[1]*ENSTA Bretagne/Lab-STICC (UMR CNRS 6285), 2 rue François Verny, 29806 Brest cedex 9, France*

✉ *E-mail: clement.roussel@ensta-bretagne.org*

Abstract: When the sea surface is illuminated by a radar sufficiently far away and in the absence of a coherent target, the scattered electromagnetic signal, called 'sea clutter' is unpredictable and can be represented by a stochastic process. A model based on stochastic differential equations and consistent with previous statistical models (K distribution) has been proposed. It depends on three parameters: $\mathscr{A}$, $\mathscr{B}$, and α. To estimate them, maximum likelihood estimators for $\mathscr{A}$ and $\mathscr{B}$, and an estimator based on the hypothesis of ergodicity for α are proposed in this study. The authors compare three expressions for the transition probabilities: the exact one, Euler's approximation, and Nowman's approximation. By regenerating the trajectories from the same Brownian increments, they can quantify the typical error made on the sea clutter from the typical error made on the estimated parameters. Though the exact transition probabilities minimise the error on the sea clutter, they show that an approximation such as Euler's is sufficient.

1 Introduction

A radar illuminating the sea surface from an airplane or a satellite has a ground resolution cell ranging from better than a metre for aircraft systems, metres for satellite systems to an upper limit of hundreds of metres (depending on whether the real or synthetic aperture is considered) [1]. The value of the sea clutter integrates all the reflectors within the cell from the capillary waves (centimetres) to the swell if there is (up to hundreds of metres). Since it is impossible to predict the time evolution of the sea surface height at the centimetre level over an area that is large, no deterministic computation of the scattered electromagnetic signal can be carried out. If the resolution cell contains enough reflectors, the usual alternative approach is to use statistical models (e.g. the K distribution), in which the sea clutter at any time t is modelled as a random variable X_t (see [2]). A theoretical limitation of this approach is that the exact relation between the sea clutters at two subsequent times, X_{t_1} and X_{t_2}, is usually not modelled. This model is, therefore, unsatisfactory to express the sea clutter as a random process, i.e. as a structured time series of random variables in opposition to a mere juxtaposition of unrelated random variables. The model proposed by Field [3] offers a solution to this problem under some hypotheses. It starts from the fundamental idea that the received signal from any random medium is a sum of contributions from dynamic scatterers of a random population. By taking the limit for very large populations of weak scatterers, Field demonstrates that the sea clutter, as a random process, is the solution to stochastic differential equations (SDEs; see Section 2). An SDE expresses a relation between the infinitesimal increment of a stochastic process between times t and $t + \mathrm{d}t$ and the value of the stochastic process at time t. Consequently, the solution of the SDE is a Markov stochastic process (depending upon the most recent past only). Therefore, in Field's model, the sea clutter can only be Markovian, which is a restriction. Although this is a debatable assumption, it is accepted as such in this study. Note that to relax the Markov assumption in the context of SDE, one could try to model the sea clutter with diffusion type SDE, in which the increment between times t and $t + \mathrm{d}t$ can depend on the whole past of the process (see chapter 6 of [4] for a proper definition). To the best of our knowledge, such a model for the sea clutter does not exist.

It has been argued in [3, 5] that Field's model could lead naturally to various applications, e.g. in radar imagery or in anomaly detection (due to the presence of a target). This model depends on three parameters: $\mathscr{A}$, $\mathscr{B}$, and α. They depend on both the sea state and the emitted electromagnetic wave (frequency and polarisation).

The issue of estimating these parameters has not been addressed so far. Many different methods for estimating the parameters of SDE exist in the literature [maximum likelihood (ML), method of moments etc.] [6]. The ML estimator has been defended as the preferred choice of estimator due to its asymptotic properties (e.g. [7, 8]). Even if we restrain to this method, there is a large freedom as for how the estimation should be made. The first degree of freedom relates to the computation of the transition probabilities. The best option is to use analytical expressions if they are available. A more simplistic approach is to use the Euler approximation, in which case the transition probability follows a Gaussian law from which it can be very straightforward to maximise the likelihood with respect to the parameter. Between the analytical formula and the Euler approximation, several degrees of complexity exist to approximate the transition probabilities (see [9]). The following question must be asked: is it always the most sophisticated approach the best choice? Furthermore, a method might work if the input time series is long and fine enough, which constitutes a second degree of freedom. However, the application behind the mathematical model sets constraints on the length and time step of a time series that could be recorded in the real world. A last implicit question remains: what are my criteria to state that the estimation is satisfactory or not?

In this study, we address these questions. We develop a method for estimating the parameters $\mathscr{A}$, $\mathscr{B}$, and α, and confront its ability to estimate correctly in relation to the specificities of our application. Our approach is based on the ML estimator and ergodicity and relies largely on numerical simulations of the SDE.

In Section 2, we explicit the stochastic model of the sea clutter and briefly explain the theoretics behind it. In Section 3, we generate trajectories with known parameters and try to retrieve the parameters by ML estimation. We propose and compare three methods for estimating $\mathscr{A}$ and $\mathscr{B}$: Euler approximation, Nowman's approximation, and the exact closed-form expressions for the transition probabilities. α is estimated differently using the ergodicity property of the sea clutter. Then, a series of numerical simulations are carried out in order to validate and compare these approaches. Section 4 is a discussion, in which we compare trajectories generated with true parameters and estimated parameters. Depending on the estimation method (Euler, Nowman or exact transition probabilities), we get a typical error made on the sea clutter and use it to select a method in several different configurations. Section 5 concludes.

2 Theoretical background

2.1 Field's model of the reflectivity

The reflectivity of the sea surface is defined as the ratio of the reflected wave to the incoming wave. At the surface of the sea, this relation can be written as

$$E_s(t) = C\Psi(t)E_i(t), \tag{1}$$

where $E_s(t)$ and $E_i(t)$ are, respectively, the scattered and incident waves (one component of the electric or magnetic wave). $C\Psi(t)$ is the complex reflectivity. C is a positive constant equal to the square root of the mean radar cross-section. $\Psi(t)$, denoted as Ψ_t in stochastic process theory, is the normalised complex reflectivity (simply referred to as reflectivity). Actually, Ψ_t modulates both the amplitude and phase of the wave. It has been demonstrated in ([3], chapters 6–8) that the reflectivity can be expressed by:

$$\Psi_t = x_t^{1/2}\gamma_t = x_t^{1/2}\left(\gamma_t^{(\mathrm{R})} + \mathrm{i}\gamma_t^{(\mathrm{I})}\right), \tag{2}$$

where x_t is the radar cross-section (RCS) normalised by its mean value. In Field's model, the RCS x_t is the total number of scatterers relative to a reference value and it is noteworthy that x_t is a stochastic process, governed by a birth–death–immigration population model. It is assumed that the scatterers have the same amplitude and are independent. γ_t is the complex-valued speckle (with unit mean amplitude), expressed with its real and imaginary parts $\gamma_t^{(\mathrm{R})}$ and $\gamma_t^{(\mathrm{I})}$. x_t, $\gamma_t^{(\mathrm{R})}$ and $\gamma_t^{(\mathrm{I})}$ are stochastic process solutions to the following SDEs ([3], chapter 8):

$$\begin{cases} \mathrm{d}x_t = \mathscr{A}(1 - x_t)\mathrm{d}t + \left(2\dfrac{\mathscr{A}}{\alpha}x_t\right)^{\frac{1}{2}}\mathrm{d}W_t^{(x)}, \\ \mathrm{d}\gamma_t^{(\mathrm{R})} = -\dfrac{1}{2}\mathscr{B}\gamma_t^{(\mathrm{R})}\mathrm{d}t + \dfrac{1}{\sqrt{2}}\mathscr{B}^{\frac{1}{2}}\mathrm{d}W_t^{(R)}, \\ \mathrm{d}\gamma_t^{(\mathrm{I})} = -\dfrac{1}{2}\mathscr{B}\gamma_t^{(\mathrm{I})}\mathrm{d}t + \dfrac{1}{\sqrt{2}}\mathscr{B}^{\frac{1}{2}}\mathrm{d}W_t^{(\mathrm{I})}, \end{cases} \tag{3}$$

where $W_t^{(x)}, W_t^{(\mathrm{R})}, W_t^{(\mathrm{I})}$ are three independent Brownian motions. It should be noted that $\gamma_t^{(\mathrm{R})}$ and $\gamma_t^{(\mathrm{I})}$ are assumed to be independent. The first equation of (3) is equation (8.9) of [3] (in [3] αx_t has been replaced by x_t) while the last two equations of (3) constitute the complex-valued equation (8.4) of [3].

Three constants parameterise the model: $\mathscr{A}$ and α for the RCS, and $\mathscr{B}$ for the speckle. $\mathscr{A}$ and α are from the underlying birth–death–immigration population model for the number of scatterers. $\mathscr{A}$ and $\mathscr{B}$ are homogeneous to the inverse of a time (i.e. a frequency). $\mathscr{A}$ can be understood as the inverse of a decorrelation time for the RCS, and $\mathscr{B}$ as the inverse of a decorrelation time for the speckle. α is the inverse of the variance of the RCS x_t for any time t as seen below.

x_t, $\gamma_t^{(\mathrm{R})}$ and $\gamma_t^{(\mathrm{I})}$ are stationary processes in Field's theory. They have the following probability densities valid for all t (see [3, 5]):

$$\begin{cases} p(x_t = x) = \dfrac{\alpha^\alpha x^{\alpha-1}\mathrm{e}^{-\alpha x}}{\Gamma(\alpha)}, \\ p\left(\gamma_t^{(\mathrm{R})} = x\right) = p\left(\gamma_t^{(\mathrm{I})} = x\right) = \dfrac{1}{\sqrt{\pi}}\mathrm{e}^{-x^2}, \end{cases} \tag{4}$$

which do not depend on $\mathscr{A}$ nor $\mathscr{B}$.

2.2 K-distribution

In this subsection, we explicit the connection between Field's model [3] and the usual K-distribution (see [2]) and observe that Field's model is in agreement with the K-distribution. Indeed, let

$$C^2 z_t = \left|C\Psi_t\right|^2 = C^2 x_t(\gamma_t^{(\mathrm{R})2} + \gamma_t^{(\mathrm{I})2}) \tag{5}$$

be the intensity, with $z_t = |\Psi_t|^2$. From the first equation of (4), we have

$$p(C^2 x_t = x) = \frac{\left(\alpha/C^2\right)^\alpha x^{\alpha-1}\mathrm{e}^{-(\alpha/C^2)x}}{\Gamma(\alpha)}. \tag{6}$$

This is the gamma distribution for the RCS as expressed in equation (4.24), p. 109 of [2], with $b = \alpha/C^2$ and $\nu = \alpha$, where b and ν are usually referred to as the scale and shape parameters, respectively. It is somewhat more intuitive to work with C^2 rather than b since

$$\mathbb{E}[C^2 x_t] = C^2. \tag{7}$$

From the second equation of (4)

$$\sqrt{2}\gamma_t^{(\mathrm{R})} \sim \sqrt{2}\gamma_t^{(\mathrm{I})} \sim \mathcal{N}(0, 1) \tag{8}$$

and by the independence of $\gamma_t^{(\mathrm{R})}$ and $\gamma_t^{(\mathrm{I})}$

$$\sqrt{2}\gamma_t^{(\mathrm{R})2} + \sqrt{2}\gamma_t^{(\mathrm{I})2} \sim \chi_2^2, \tag{9}$$

from which we obtain

$$\begin{aligned} &p\left(\gamma_t^{(\mathrm{R})2} + \gamma_t^{(\mathrm{I})2} = x\right) = \mathrm{e}^{-x} \\ &\Leftrightarrow p\left(u(\gamma_t^{(\mathrm{R})2} + \gamma_t^{(\mathrm{I})2}) = x\right) = \frac{\mathrm{e}^{-x/u}}{u} \end{aligned} \tag{10}$$

for all $u \geq 0$. Then

$$\begin{aligned} p(C^2 z_t = x) &= \int_0^{+\infty} p(C^2 x_t = u)p(C^2 z_t = x \big| C^2 x_t = u)\mathrm{d}u \\ &= \int_0^{+\infty} \frac{(\alpha/C^2)^\alpha x^{\alpha-1}\mathrm{e}^{-(\alpha/C^2)x}}{\Gamma(\alpha)}\frac{\mathrm{e}^{-x/u}}{u}\mathrm{d}u \\ &= \frac{2b^{(\nu+1)/2}x^{(\nu-1)/2}}{\Gamma(\nu)}K_{\nu-1}(2\sqrt{bx}). \end{aligned} \tag{11}$$

We have retrieved the K-distribution for the intensity (see equation (4.26), p. 109 of [2]).

3 Estimation of the SDE parameters

In this section, we propose to estimate $\mathscr{A}$ and $\mathscr{B}$ by ML and α using ergodicity. From the radar data, one can already note that the speckle variation timescale is 10 ms while the RCS variation timescale is about 1 s [10, 11]. Therefore, the corresponding orders of magnitude of $\mathscr{A}$ and $\mathscr{B}$ are $\mathscr{A} = 1$ Hz and $\mathscr{B} = 100$ Hz. α is dimensionless and its value typically ranges from 0.1 to $+\infty$ ([2], p. 110–111).

In a real data situation, we would have access to the complex-valued time series $C\tilde{\Psi}_t$. C^2 is the mean sea clutter power, and we shall refer the reader to chapter 9 of [2] for its estimation. In this study, since we focus on the estimation of the SDE parameters, we set $C = 1$. Using the difference between the slow dynamics of x_t and fast dynamics of $\gamma_t^{(\mathrm{R})}$, $\gamma_t^{(\mathrm{I})}$, one can retrieve x_t (see [12]) and then $\gamma_t^{(\mathrm{R})}$ and $\gamma_t^{(\mathrm{I})}$ from the observation of Ψ_t. We, therefore, assume in Sections 3.1 and 3.2 (estimation of $\mathscr{A}$, $\mathscr{B}$, and α) that we observe three discrete time series: $\tilde{x}_i, \tilde{\gamma}_i^{(\mathrm{R})}, \tilde{\gamma}_i^{(\mathrm{I})}$, where the measurements have been made at times t_i with i ranging from 0 to n. We also assume for simplicity that $\forall i,\ t_i - t_{i-1} = \Delta t$. We notice that $\gamma_t^{(\mathrm{R})}$ and $\gamma_t^{(\mathrm{I})}$ follow the same SDE. Since they are independent, it is sufficient to use $\tilde{\gamma}_i^{(\mathrm{R})}$ to estimate $\mathscr{B}$. Alternatively, we can average the two estimates arising from these two time series. A realistic value for Δt is 0.001 s since the pulse repetition frequency is about 1 kHz for satellite and airborne applications [1].

3.1 ML estimation of $\mathscr{A}$ and $\mathscr{B}$

The ML estimation consists of maximising the likelihood function with respect to the parameters, i.e. determining the parameters that make the observed data the most probable. In the present case, the likelihood function of x_t depends on $\mathscr{A}$ and α as described in the stochastic differential model equation (3). Since x_t is a Markov process (it is a solution to a SDE, see [13]), the joint probability density function associated with process x_t is

$$\begin{aligned} &p(x_{t_0} = \tilde{x}_0, \ldots, x_{t_n} = \tilde{x}_n) \\ &\quad = p(x_{t_0} = \tilde{x}_0, \ldots, x_{t_{n-1}} = \tilde{x}_{n-1}) \\ &\qquad \times p(x_{t_n} = \tilde{x}_n \big| x_{t_0} = \tilde{x}_0, \ldots, x_{t_{n-1}} = \tilde{x}_{n-1}) \\ &\quad = p(x_{t_0} = \tilde{x}_0, \ldots, x_{t_{n-1}} = \tilde{x}_{n-1}) \\ &\qquad \times p(x_{t_n} = \tilde{x}_n \big| x_{t_{n-1}} = \tilde{x}_{n-1}) . \end{aligned}$$

Thus, the likelihood function for the RCS is

$$\mathscr{L}(\mathscr{A}, \alpha) = p(x_{t_0} = \tilde{x}_0) \prod_{i=1}^{n} p(x_{t_i} = \tilde{x}_i | x_{t_{i-1}} = \tilde{x}_{i-1}) . \tag{12}$$

Similarly, the likelihood function for the speckle is

$$\mathscr{L}(\mathscr{B}) = p\left(\gamma_{t_0}^{(\mathrm{R})} = \tilde{\gamma}_0^{(\mathrm{R})}\right) \prod_{i=1}^{n} p\left(\gamma_{t_i}^{(\mathrm{R})} = \tilde{\gamma}_i^{(\mathrm{R})} \Big| \gamma_{t_{i-1}}^{(\mathrm{R})} = \tilde{\gamma}_{i-1}^{(\mathrm{R})}\right). \tag{13}$$

In this section, we compare three different models for the transition probabilities of x_t and $\gamma_t^{(\mathrm{R})}$: the Euler approximation, Nowman's approximation, and the exact transition probabilities.

3.1.1 Euler's approximation: Euler's approximation relies on the discretisation method of Euler–Maruyama, which enables both to simulate numerical trajectories (i.e. solve numerically the SDE) and to have an analytical discrete-time approximation of our continuous-time process. In Euler's approximation, the drift and volatility of the SDE are assumed constant over the interval $[t_{i-1}, t_i]$ (see the Appendix). Applied to the SDE of x_t, i.e. the first equation of (3), and assuming a constant time step Δt, Euler–Maruyama's scheme gives

$$x_{t_i} \simeq x_{t_{i-1}} + \mathscr{A}(1 - x_{t_{i-1}})\Delta t + \left(\frac{2\mathscr{A}}{\alpha} x_{t_{i-1}}\right)^{1/2} \mathscr{N}(0, \Delta t) . \tag{14}$$

Applying Euler–Maruyama's scheme to the second equation of (3), we get

$$\gamma_{t_i}^{(\mathrm{R})} \simeq \gamma_{t_{i-1}}^{(\mathrm{R})} - \frac{1}{2}\mathscr{B}\gamma_{t_{i-1}}^{(\mathrm{R})}\Delta t + \frac{1}{\sqrt{2}}\mathscr{B}^{1/2}\mathscr{N}(0, \Delta t) . \tag{15}$$

From (14) and (15), we get the following transition probabilities:

$$\begin{cases} p(x_{t_i} = \tilde{x}_i \big| x_{t_{i-1}} = \tilde{x}_{i-1}) = \dfrac{\sqrt{\alpha}}{\sqrt{4\pi\tilde{x}_{i-1}\mathscr{A}\Delta t}} \mathrm{e}^{-\frac{\alpha(\tilde{x}_i - \mathscr{A}\Delta t - (1 - \mathscr{A}\Delta t)\tilde{x}_{i-1})^2}{4\mathscr{A}\Delta t\tilde{x}_{i-1}}}, \\ p\left(\gamma_{t_i}^{(\mathrm{R})} = \tilde{\gamma}_i^{(\mathrm{R})} \Big| \gamma_{t_{i-1}}^{(\mathrm{R})} = \tilde{\gamma}_{i-1}^{(\mathrm{R})}\right) = \dfrac{1}{\sqrt{\pi\mathscr{B}\Delta t}} \mathrm{e}^{-\frac{\left(\gamma_i^{(\mathrm{R})} - \tilde{\gamma}_{i-1}^{(\mathrm{R})}(1 - \mathscr{B}\Delta t/2)\right)^2}{\mathscr{B}\Delta t}}, \end{cases} \tag{16}$$

We inject these expressions in the corresponding likelihood functions (12) and (13). Using (12) and assuming that we know α, we express the condition $\partial \ln \mathscr{L} / \partial \mathscr{A}(\tilde{\mathscr{A}}, \alpha) = 0$ to estimate $\mathscr{A}$. We also express the condition $\partial \ln \mathscr{L} / \partial \mathscr{B}(\tilde{\mathscr{B}}) = 0$. $\tilde{\mathscr{A}}$ and $\tilde{\mathscr{B}}$ are the values of $\mathscr{A}$ and $\mathscr{B}$, which maximise the likelihood. After some calculations, it is possible to show that $\tilde{\mathscr{A}}$ and $\tilde{\mathscr{B}}$ are the roots of two second-order polynomials

$$\begin{cases} -\sum_{i=1}^{n} \dfrac{\alpha(\tilde{x}_{i-1} - 1)^2}{4\tilde{x}_{i-1}} \mathscr{A}^2 - \dfrac{n}{2}\mathscr{A} + \sum_{i=1}^{n} \dfrac{\alpha(\tilde{x}_i - \tilde{x}_{i-1})^2}{4\tilde{x}_{i-1}\Delta t} = 0, \\ -\sum_{i=1}^{n} \dfrac{\tilde{\gamma}_{i-1}^{(\mathrm{R})2}\Delta t}{4} \mathscr{B}^2 - \dfrac{n}{2}\mathscr{B} + \sum_{i=1}^{n} \dfrac{\left(\tilde{\gamma}_i^{(\mathrm{R})} - \tilde{\gamma}_{i-1}^{(\mathrm{R})}\right)^2}{\Delta t} = 0. \end{cases} \tag{17}$$

For both $\mathscr{A}$ and $\mathscr{B}$, the discriminant of the polynomial is always positive and there is only one positive root (no ambiguity). $\tilde{\mathscr{A}}$ and $\tilde{\mathscr{B}}$ are therefore easily found analytically without numerical minimisation of the likelihood function.

3.1.2 Nowman's approximation: Nowman's approach [9, 14] is applicable to SDE with a linear drift, i.e. of the form

$$\mathrm{d}X_t = \kappa(\mu - X(t))\mathrm{d}t + \sigma(X_t)\mathrm{d}W_t, \tag{18}$$

where κ and μ are two constants. This is precisely the case for the SDE of x_t and $\gamma_t^{(\mathrm{R})}$, hence our interest in this approach. In the case of a linear drift, the solution to the SDE (18) can be expressed as follows:

$$\begin{aligned} X_{t_i} = X_{t_{i-1}} \mathrm{e}^{-\kappa\Delta t} + \mu(1 - \mathrm{e}^{-\kappa\Delta t}) \\ + \int_{t_{i-1}}^{t_i} \sigma(X_s) \mathrm{e}^{-\kappa(t_i - s)} \mathrm{d}W_s . \end{aligned} \tag{19}$$

Nowman assumes that the volatility is constant over the time interval: $\forall s \in [t_{i-1}, t_i]\ \sigma(X_s) = \sigma(X_{t_{i-1}})$, in which case

$$X_{t_i} = X_{t_{i-1}} \mathrm{e}^{-\kappa\Delta t} + \mu(1 - \mathrm{e}^{-\kappa\Delta t}) + \eta_{t_i} \tag{20}$$

with

$$\eta_{t_i} = \sigma(X_{t_{i-1}}) \int_{t_{i-1}}^{t_i} \mathrm{e}^{-\kappa(t_i - s)} \mathrm{d}W_s . \tag{21}$$

We can show that $\mathbb{E}(\eta_{t_i}) = 0$ and $\mathbb{E}(\eta_{t_i}^2) = \sigma(X_{t_{i-1}})^2\left(1 - \mathrm{e}^{-2\kappa\Delta t}\right)/2\kappa$ from which we make the approximation

$$\eta_{t_i} \sim \mathscr{N}\left(0, \frac{\sigma(X_{t_{i-1}})^2\left(1 - \mathrm{e}^{-2\kappa\Delta t}\right)}{2\kappa}\right). \tag{22}$$

Combining (20) and (22), we get

$$\begin{aligned} X_{t_i} = X_{t_{i-1}} \mathrm{e}^{-\kappa\Delta t} + \mu(1 - \mathrm{e}^{-\kappa\Delta t}) \\ + \mathscr{N}\left(0, \frac{\sigma(X_{t_{i-1}})^2\left(1 - \mathrm{e}^{-2\kappa\Delta t}\right)}{2\kappa}\right). \end{aligned} \tag{23}$$

The transition probabilities are again Gaussian but the drift has been completely resolved. For x_t and $\gamma_t^{(\mathrm{R})}$, we get (see (24)) . Then, we inject these expressions in the corresponding likelihood functions (12) and (13). The conditions $\partial \ln \mathscr{L} / \partial \mathscr{A}(\tilde{\mathscr{A}}, \alpha) = 0$ and

$$\begin{cases} p(x_{t_i} = \tilde{x}_i \big| x_{t_{i-1}} = \tilde{x}_{i-1}) = \dfrac{\sqrt{\alpha}}{\sqrt{2\pi\tilde{x}_{i-1}(1 - \mathrm{e}^{-2\mathscr{A}\Delta t})}} \mathrm{e}^{-\frac{1}{2}\frac{\alpha\left(\tilde{x}_i - 1 + \mathrm{e}^{-\mathscr{A}\Delta t}(1 - \tilde{x}_{i-1})\right)^2}{\tilde{x}_{i-1}(1 - \mathrm{e}^{-2\mathscr{A}\Delta t})}}, \\ p\left(\gamma_{t_i}^{(\mathrm{R})} = \tilde{\gamma}_i^{(\mathrm{R})} \Big| \gamma_{t_{i-1}}^{(\mathrm{R})} = \tilde{\gamma}_{i-1}^{(\mathrm{R})}\right) = \dfrac{1}{\sqrt{\pi(1 - \mathrm{e}^{-\mathscr{B}\Delta t})}} \mathrm{e}^{-\frac{\left(\tilde{\gamma}_i^{(\mathrm{R})} - \tilde{\gamma}_{i-1}^{(\mathrm{R})}\mathrm{e}^{-\mathscr{B}\Delta t/2}\right)^2}{1 - \mathrm{e}^{-\mathscr{B}\Delta t}}}. \end{cases} \tag{24}$$

$\partial \ln \mathscr{L}/\partial \mathscr{B}(\tilde{\mathscr{B}}) = 0$ lead to third-order polynomials depending on $X = \mathrm{e}^{-\tilde{\mathscr{A}}\Delta t}$ and $Y = \mathrm{e}^{-\tilde{\mathscr{B}}\Delta t/2}$

$$\begin{cases} nX^3 - \sum_{i=1}^{n} \dfrac{\alpha(\tilde{x}_i - 1)(\tilde{x}_{i-1} - 1)}{\tilde{x}_{i-1}} X^2 \\ \quad + \left(-n + \sum_{i=1}^{n} \dfrac{\alpha(\tilde{x}_i - 1)^2 + \alpha(\tilde{x}_{i-1} - 1)^2}{\tilde{x}_{i-1}} \right) X \\ \quad - \sum_{i=1}^{n} \dfrac{\alpha(\tilde{x}_i - 1)(\tilde{x}_{i-1} - 1)}{\tilde{x}_{i-1}} = 0, \\ \dfrac{n}{2} Y^3 - \sum_{i=1}^{n} \tilde{\gamma}_{i-1}^{(R)} \tilde{\gamma}_i^{(R)} Y^2 \\ \quad + \left(-\dfrac{n}{2} + \sum_{i=1}^{n} \left(\tilde{\gamma}_i^{(R)2} + \tilde{\gamma}_{i-1}^{(R)2} \right) \right) Y \\ \quad - \sum_{i=1}^{n} \tilde{\gamma}_{i-1}^{(R)} \tilde{\gamma}_i^{(R)} = 0. \end{cases} \tag{25}$$

There is only one real root for each of these two polynomials (the other two being complex conjugates). Thus $\tilde{\mathscr{A}}$ and $\tilde{\mathscr{B}}$ are determined analytically neither with ambiguity nor with the need for numerical minimisation. In this regard, Nowman's approximation is very similar to Euler's approximation, but it remains formally better since the drift is exactly solved.

3.1.3 Exact transition probabilities: The transition probabilities of x_t and $\gamma_t^{(\mathrm{R})}$ can be obtained by solving their respective Fokker–Planck equations (see [15]). $\gamma_t^{(\mathrm{R})}$ is an Ornstein–Uhlenbeck process and its Fokker–Planck equation is easily solved by Fourier transform and the method of characteristics ([15], chapter 5). In the particular case of $\gamma_t^{(\mathrm{R})}$, it gives [5]

$$p\left(\gamma_{t+\Delta t}^{(\mathrm{R})} = x \middle| \gamma_t^{(\mathrm{R})} = y\right) = \frac{1}{\sqrt{2\pi v(\Delta t)}} \mathrm{e}^{-\frac{1}{2}\frac{(x - m(\Delta t))^2}{v(\Delta t)}}, \tag{26}$$

which is Gaussian with the expectation

$$m(\Delta t) = y\mathrm{e}^{-\mathscr{B}\Delta t/2} \tag{27}$$

and variance

$$v(\Delta t) = \frac{1 - \mathrm{e}^{-\mathscr{B}\Delta t}}{2}. \tag{28}$$

Replacing x, y, t, $t + \Delta t$ by $\tilde{\gamma}_i^{(\mathrm{R})}$, $\tilde{\gamma}_{i-1}^{(\mathrm{R})}$, t_{i-1}, t_i, respectively, we obtain the same expression as the second equation of (24). This is not surprising since Nowman's method resolves the drift and then assumes locally a constant volatility. From (3), we see that the constant volatility assumption is actually always true for $\gamma_t^{(\mathrm{R})}$.

x_t is a Cox–Ingersoll–Ross process. It has been shown that its transition probabilities can be written analytically ([3, p. 63, 16])

$$p(x_{t+\Delta t} = x | x_t = y) = c\mathrm{e}^{-cx - c\delta y} \left(\frac{x}{y\delta} \right)^{(\alpha-1)/2} I_{\alpha-1}(2c\sqrt{xy\delta}), \tag{29}$$

where $c = \alpha/1 - \mathrm{e}^{-\mathscr{A}\Delta t}$, $\delta = \mathrm{e}^{-\mathscr{A}\Delta t}$ and $I_{\alpha-1}$ is the modified Bessel function of the first kind of order $\alpha - 1$.

Maximisation of the likelihood function is achieved numerically by gradient descent applied to $-\ln(\mathscr{L}(\mathscr{A}, \alpha))$, α being fixed.

3.1.4 Estimation of $\mathscr{A}$ and $\mathscr{B}$ from simulated trajectories: To assess the ability of a method to estimate the parameters, we simulate many trajectories of x_t and $\gamma_t^{(\mathrm{R})}$ with known parameters and then try to retrieve them. To do so, we solve numerically the SDE in (3) using Euler–Maruyama's scheme for $\gamma_t^{(\mathrm{R})}$, $\gamma_t^{(\mathrm{I})}$ and Milstein's scheme for x_t (see the Appendix and [17] for an introduction to numerical simulation of the SDE). The Euler–Maruyama scheme is not used for x_t because it raises numerical issues. Indeed, it could generate negative values, which is absurd since the RCS is always positive.

In our experiments, the simulation time step, $\hat{\Delta}t$, should be much smaller than the measurement time step Δt quoted above. Otherwise, the numerical trajectory will not replicate adequately the dynamics of the original SDE, with its time-dependent drifts and volatilities, and the transition probabilities of the simulated trajectories will be very different from the real ones. Consequently, the generated trajectories are evaluated at times $\hat{t}_k$ for $k = 0, \ldots, mn$, where m is the decimation ratio. Then, for the estimation stage, they are down-sampled (i.e. decimated) to the times t_i for $i = 0, \ldots, n$, with $t_i = \hat{t}_{mi}$. In our numerical simulations, Δt being fixed at 0.001 s, we choose $\hat{\Delta}t = 10^{-5}$ s leading to a decimation ratio $m = 100$. Note that we have verified that $\hat{\Delta}t = 10^{-5}$ s is small enough: changing it to $\hat{\Delta}t = 10^{-7}$ s does not alter the results presented here but it requires a much longer computing time.

To estimate $\mathscr{A}$ and $\mathscr{B}$ (with the order of magnitude around 1 and 100 Hz, respectively), a duration of 1 s is sufficient leading to $t_n = t_{1000} = 1$ s. The estimations of $\mathscr{A}$ and $\mathscr{B}$ are completely independent since $\mathscr{A}$ requires only the trajectory of the RCS and $\mathscr{B}$ requires only the trajectory of the speckle. However, the estimation of $\mathscr{A}$ does require α, and for simplicity, we set its value to 1. We will see afterwards that α can be estimated easily and in a meaningful way, with the constraint that the time series of the RCS for its estimation should be much longer than 1 s.

3.1.5 Estimation of $\mathscr{A}$: For the proposed numerical simulations, the true values of $\mathscr{A}$ range in the interval [0.1, 10] Hz with a step of 0.1 Hz. For each value of $\mathscr{A}$, we generate $N = 1000$ trajectories of the RCS (with $\alpha = 1$) of duration 1 s and time step 10^{-5} s. As explained in the previous section, the N estimates $\tilde{\mathscr{A}}_1, \tilde{\mathscr{A}}_2, \ldots, \tilde{\mathscr{A}}_N$ are obtained using the down-sampled trajectories (i.e. trajectories with time step equal to 0.001 s). Then, the estimation bias $b(\mathscr{A})$ and (unbiased) standard deviation $\sigma(\mathscr{A})$ are computed

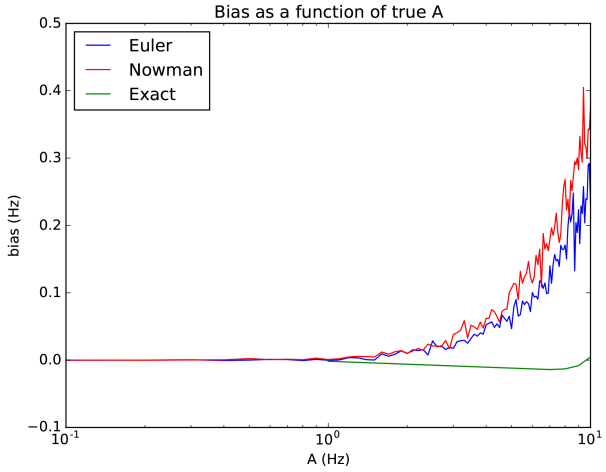
$$b(\mathscr{A}) = \frac{1}{N} \sum_{i=1}^{N} (\tilde{\mathscr{A}}_i - \mathscr{A}), \tag{30}$$

$$\sigma^2(\mathscr{A}) = \frac{1}{N-1} \sum_{i=1}^{N} (\tilde{\mathscr{A}}_i - \mathscr{A})^2. \tag{31}$$

Fig. 1 presents the estimation bias and standard deviations in this configuration, as a function of the true value of $\mathscr{A}$ used to generate the trajectories. The Euler and Nowman approximations have very similar results, with no improvement when using Nowman's approximation. On the contrary, a slight increase in bias is observed compared to Euler's approximation. For both of them, there is an exponential increase in the bias and standard deviation, which persist if we compute the relative bias and standard deviation (i.e. $b(\mathscr{A})/\mathscr{A}$ and $\sigma(\mathscr{A})/\mathscr{A}$). However, using the exact transition probabilities completely annihilates the bias. Even though the standard deviation still increases exponentially, this is only due to the log scale. The relative variance is constant.

3.1.6 Estimation of $\mathscr{B}$: For the numerical simulation in this part, the true values of $\mathscr{B}$ range in the interval [10, 1000] Hz with a step of 10 Hz. Like in the previous section, for each $\mathscr{B}$, we generate $N = 1000$ trajectories of $\gamma_t^{(\mathrm{R})}$ of duration 1 s and time step 10^{-5} s. Then, we obtain N estimates $\tilde{\mathscr{B}}_1, \tilde{\mathscr{B}}_2, \ldots, \tilde{\mathscr{B}}_N$ from the down-sampled trajectories. Finally, the estimation bias $b(\mathscr{B})$ and standard deviation $\sigma(\mathscr{B})$ are then computed.

Fig. 2 presents the estimation bias and standard deviations in this configuration, as a function of the true value of $\mathscr{B}$ used to generate the trajectories. The Euler approximation results in an



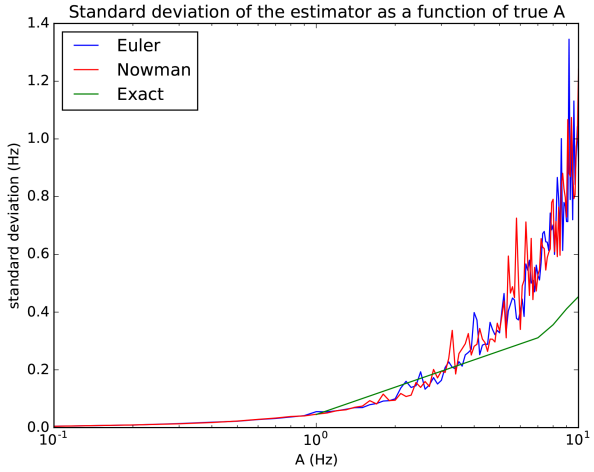


Fig. 1 *Estimation bias (up) and standard deviation (down) of $\mathscr{A}$ as a function of true $\mathscr{A}$ based on 1000 trajectories of duration 1 s each. Three methods are compared: the Euler approximation, the Nowman approximation, and the exact transition probability*

exponential increase of the absolute value of the bias and standard deviation with increasing $\mathscr{B}$, which holds for the relative bias and standard deviation. However, using the Nowman approximation, which is the exact transition probabilities for this case, completely annihilates the bias. Even though the standard deviation still increases exponentially, this is mostly due to the log scale. The relative variance increases only slightly with increasing $\mathscr{B}$. We notice that the standard deviation is lower for Euler's approximation, especially at high value of $\mathscr{B}$. We would expect the otherwise since the exact transition probabilities are theoretically more relevant than Euler's approximation. However, $\mathscr{B} = 1000$ Hz is rather extreme relative to the reference value $\mathscr{B} = 100$ Hz quoted before. From 10 Hz up to 350 Hz, the relative difference of the standard deviation between Euler and Nowman is below 1 %.

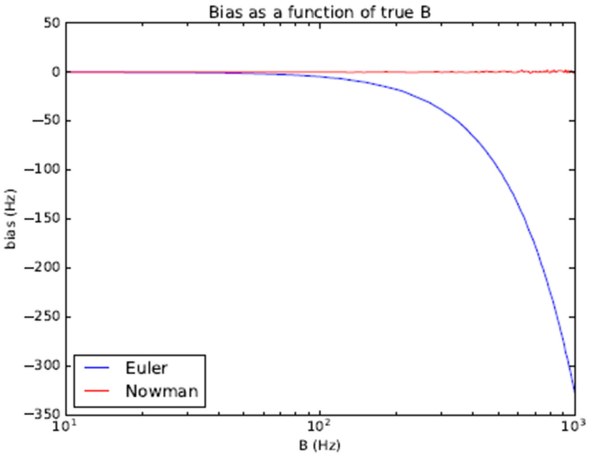
3.2 Estimation of α

The estimation of α is much simpler than that of $\mathscr{A}$ and $\mathscr{B}$. We assume that the sea clutter is ergodic and use this property to estimate the parameter. This assumption yields a simple estimator which has the disadvantage of requiring longer trajectories, as will be observed thereafter. It arises as a necessity to have sufficient decorrelated samples in the time series, although strictly speaking it is never the case that two samples are decorrelated.

According to (4), $\forall t$

$$\mathrm{var}(x_t) = \frac{1}{\alpha}. \tag{32}$$

We assume that the RCS is ergodic, or more precisely we assume that



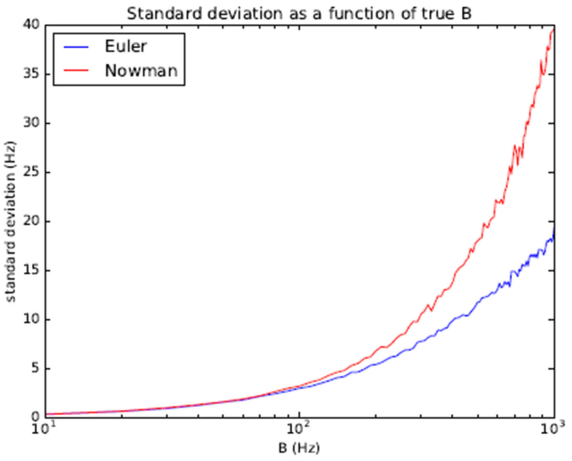


Fig. 2 *Estimation bias (up) and standard deviation (down) of $\mathscr{B}$ as a function of true $\mathscr{B}$ based on 1000 trajectories of duration 1 s each. Two methods are compared: the Euler approximation and the Nowman approximation, which is equal to the exact transition probability in this particular case*

$$\lim_{T \to +\infty} \frac{1}{T} \int_0^T (\tilde{x}_t - 1)^2 \mathrm{d}t = \mathbb{E}[(x_0 - 1)^2] = \frac{1}{\alpha}, \tag{33}$$

where $\tilde{x}_t$ is a particular trajectory of the process x_t. We note that $\forall t$, $\mathbb{E}[x_t] = 1$. To compute α from a time series $\tilde{x}_i$ with i ranging from 0 to n, we make the approximation

$$\begin{aligned} \frac{1}{\alpha} &= \lim_{T \to +\infty} \frac{1}{T} \int_0^T (\tilde{x}_t - 1)^2 \mathrm{d}t \simeq \frac{1}{t_n} \int_0^{t_n} (\tilde{x}_t - 1)^2 \mathrm{d}t \\ &\simeq \frac{1}{n\Delta t} \sum_{i=0}^{n-1} (\tilde{x}_i - 1)^2 \Delta t \,. \end{aligned} \tag{34}$$

Fig. 3 presents the estimation bias and standard deviation of α for 1000 trajectories. We see that even with 300 s, the standard deviation is substantial compared to the true value of α. The estimation of α requires a much longer observation of the sea surface than the estimation of $\mathscr{A}$, due to its different physical meaning and mathematical expression. $\mathscr{A}$ is the inverse of a decorrelation time and therefore a time series $1/\mathscr{A}$ is long (order of magnitude) is sufficient to have a satisfying estimation of $\mathscr{A}$. α is the variance of x_t for any t. To estimate it satisfactorily, one ought to average enough independent realisations of x_t, which is approximated using the ergodic hypothesis, i.e. by averaging over one trajectory. However, two values $\tilde{x}_{t_1}$ and $\tilde{x}_{t_2}$ of one trajectory constitute roughly two independent realisations of the same random variable only if $t_2 - t_1$ is large enough. The correlation between x_{t_1} and x_{t_2} can be shown to be $\mathrm{corr}(x_{t_1}, x_{t_2}) = \mathrm{e}^{-\mathscr{A}(t_2 - t_1)}$, which gives a correlation of 0.05 if $\mathscr{A}(t_2 - t_1) = 3$, in which case

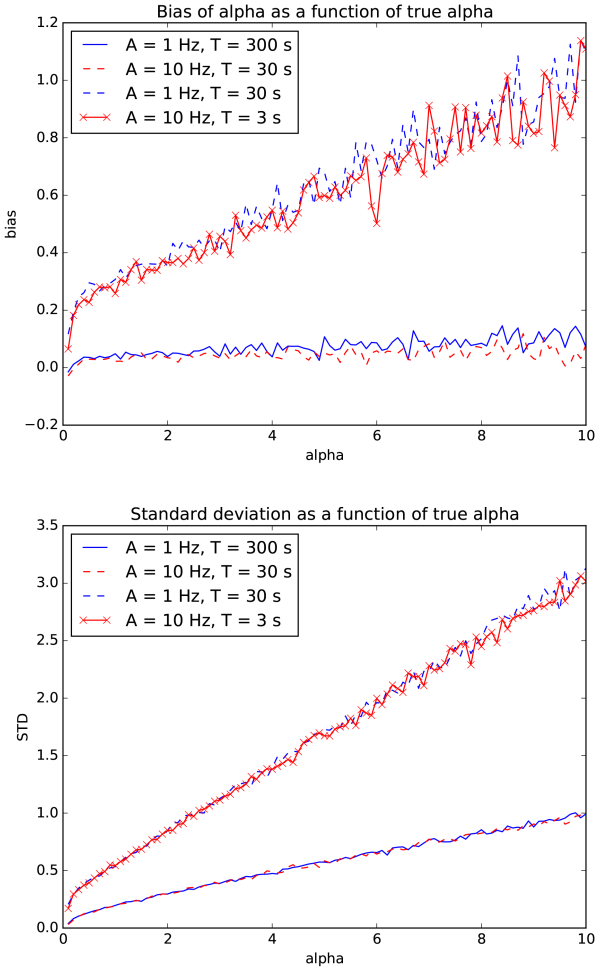


Fig. 3 *Estimation bias (up) and standard deviation (down) of α as a function of true α based on 1000 trajectories*

we can assume that x_{t_1} and x_{t_2} are independent. Fig. 3 illustrates the fact that the variance and bias of the estimators look alike for $\mathscr{A} = 1$ Hz, $T = 300$ s, and $\mathscr{A} = 10$ Hz, $T = 30$ s. This makes sense since both these configurations give ∼100 independent realisations of the same random variable. The resemblance is even more striking for $\mathscr{A} = 1$ Hz, $T = 30$ s, and $\mathscr{A} = 10$ Hz, $T = 3$ s.

3.3 Joint estimation of $\mathscr{A}$ and α

In Section 3.1, we assumed that α was known for the estimation of $\mathscr{A}$. Simultaneous estimation of $\mathscr{A}$ and α can be done using, for example, the Berndt-Hall-Hall-Hausman (BHHH) algorithm or more advanced techniques [18, 19]. We will not explore these for two reasons. The first one is that they require computing the derivative of the likelihood function (with exact transition probabilities) with respect to α, which has no closed-form expression due to the Bessel function (29). The derivative with respect to $\mathscr{A}$ is very tedious but tractable. The second reason is that α can be estimated as explained previously without any knowledge on $\mathscr{A}$. A more natural and meaningful method to estimate both α and $\mathscr{A}$ emerges.

(1) Use the ergodic assumption (34) to estimate α and get $\tilde{\alpha}$.
(2) Estimate $\mathscr{A}$ using Euler's approximation or the exact transition probabilities, with α replaced by $\tilde{\alpha}$.

We set $\alpha = 1$ and $\mathscr{A} = 1$ Hz. As previously, we generate 1000 trajectories of a fixed duration T and estimate α and $\mathscr{A}$ as explained for each of the trajectories. We focus on the estimation relative standard deviation of α and $\mathscr{A}$ as a function of the duration T, for a duration between 10 and 1000 s. We set the bias aside because it can always be corrected by indirect inference [9]. Fig. 4 presents the results of this simulation. At known α, the estimation of $\mathscr{A}$ was

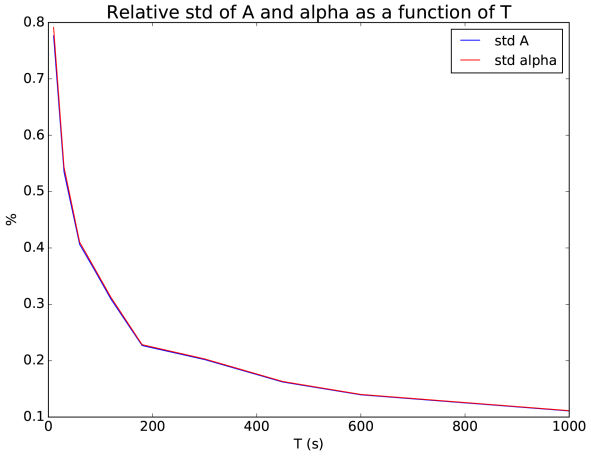


Fig. 4 *Relative standard deviations of $\mathscr{A}$ and α as a function of the duration T of the trajectories, for 1000 trajectories*

satisfactory even with 1 s long trajectories (Section 3.1.5). However, when α must also be estimated, a longer trajectory is necessary as evidenced here. If the trajectory is too short, the error on α is large, and as a consequence so is that on $\mathscr{A}$.

4 Discussion

4.1 Comparing estimators

Different estimators have different performances. For example, we saw in Section 3, that using the exact transition probabilities in the ML function is better than using Euler's approximation for at least three different reasons: it removes the bias, it reduces the variance (by a factor 2 for $\mathscr{A} = 10$ Hz, see Fig. 1), and we know a priori that it should be better. However, using the most complete model is not always necessary. We will show in this section that using Euler's approximation is sufficient to estimate $\mathscr{A}$. We use numerical simulations to convert an error made on the estimated parameter into an error made on the RCS x_t. We argue that the estimation is ultimately limited by the information carried in the trajectory that we use (sampling rate and duration).

Let us assume that an error $\delta\mathscr{A}$ is made on the estimation of $\mathscr{A}$, i.e. $\tilde{\mathscr{A}} = \mathscr{A} + \delta\mathscr{A}$. To convert it into errors on the RCS, we generate numerically trajectories of the RCS with the parameter $\mathscr{A}$ using Milstein's scheme. Its order of convergence is higher than Euler–Maruyama's but it follows the same principle: first we generate a series of Brownian increments $\Delta W_{\hat{t}_k}, k = 1, \ldots, 100n$, second, we compute the corresponding series of the RCS $\hat{x}(\mathscr{A}) = x_{\hat{t}_k}(\mathscr{A}), k = 1, \ldots, 100n$ starting from an arbitrary initial condition $x_{\hat{t}_0}$ (see the Appendix). Using the exact same Brownian increments and initial condition, we can regenerate the trajectory with the estimated parameter and obtain $\hat{x}(\mathscr{A} + \delta\mathscr{A}) = x_{\hat{t}_k}(\mathscr{A} + \delta\mathscr{A}), k = 1, \ldots, 100n$. The trajectories $\hat{x}(\mathscr{A})$ and $\hat{x}(\mathscr{A} + \delta\mathscr{A})$ are expected to be identical for small $\delta\mathscr{A}$.

Fig. 5 presents the results of this method applied to the RCS in the case where α is known and equal to 1. We generate 1000 trajectories of the RCS, $\hat{x}^{(i)}(\mathscr{A}), i = 1, \ldots, 1000$, 1 s long each and save the Brownian increments used to generate them. $\mathscr{A}$ is estimated from each trajectory, which gives $\tilde{\mathscr{A}}_i, i = 1, \ldots, 1000$. The estimation is carried out using the exact transition probabilities (unbiased estimation) and Euler's approximation. For the latter, we remove the bias by indirect inference [9]: our previous simulations already gave us the bias for $\mathscr{A} = 1$ Hz. The trajectories are regenerated with the estimated parameters, which gives $\hat{x}^{(i)}(\tilde{\mathscr{A}}_i)$, $i = 1, \ldots, 1000$. Finally, we compute the mean standard deviation between the trajectories

$$\sigma_{\text{est}} = \frac{1}{1000} \sum_{i=1}^{1000} \sigma_i \tag{35}$$

with

$$\sigma_i^2 = \frac{1}{100n-1}\sum_{k=1}^{100n}\left(x_{t_k}^{(i)}(\mathscr{A}) - x_{t_k}^{(i)}(\tilde{\mathscr{A}}_i)\right)^2. \quad (36)$$

The lower part of Fig. 5 shows the standard deviation between the original and regenerated trajectories, with a true parameter $\mathscr{A} = 1$ Hz. We plot only the first 100 values to clearly show the correlation. They are pretty much the same for the trajectories regenerated with Euler's estimation of the parameter $\mathscr{A}$, and those regenerated with the estimation of the parameter using the exact transition probabilities (see Section 3). In these simulations, the mean standard deviation is 6.9×10^{-3} for Euler's approximation and 6.77×10^{-3} for the estimation based on the exact transition probabilities. The standard deviation of the difference between the regenerated trajectory using Euler's estimation and the estimation based on exact transition probabilities (black curve) is much smaller, with an average of 9.2×10^{-4}. This can be summed up by the upper part of Fig. 5: for both estimations, the regenerated trajectories are much closer together than they are to the original trajectory. Fig. 6 confirms this: it shows that the estimated parameters (debiased for Euler's) are highly correlated when plotted as a function of trajectory number. It follows immediately that the regenerated trajectories will be very close, as well as all our comments about Fig. 5. Our interpretation is that the error on the estimation of $\mathscr{A}$ in our configuration (1 s long time series sampled at 1 ms) is mainly due to the statistical peculiarity of the trajectory, not the choice between Euler's approximation or the exact transition probabilities or maybe any other estimation method. The time series is one short sampled chunk of one time-continuous and of infinite duration realisation of the random process, and therefore carries only limited information that may be almost completely used in a method as simple as Euler's approximation.

Out of curiosity, we compared the standard deviations of Fig. 5 to the typical measurement standard deviation of a satellite radar sensor. The RCS is normally in units of a surface, generally m^2. Taking the decimal logarithm and multiplying by 10 gives the RCS expressed in dBm^2. Antony *et al.* [20] have shown that the post-calibration radiometric accuracy (or measurement standard deviation) of the TerraSAR-X and TanDEM-X satellite systems is of 0.25 dBm^2 for a target 43.5 dBm^2. By definition, x_t is the ratio of the observed sea-surface RCS (expressed as surfaces) to its mean value, we can show that the measurement standard deviation of x_t corresponding to 0.25 dBm^2 for a target of 43.5 dBm^2 is 4.73×10^{-5}, much smaller than 6.9×10^{-3} and 6.77×10^{-3}. Our estimation, regardless of the used method, is not within radiometric accuracy when converted into sea-clutter units.

4.2 Numerical schemes

Throughout this paper, we have used numerical schemes to simulate our processes. We used the Euler–Maruyama scheme for $\gamma_t^{(R)}$ and $\gamma_t^{(I)}$, and Milstein's scheme for x_t. The choice of this scheme was motivated by its simplicity. We chose Milstein's scheme for x_t only because the Euler–Maruyama scheme was evidently not good enough (generating negative values). Whether these schemes are performant enough in the context of our application can be subject to debate. However, it was observed in [5] that they lead to numerical distributions which fit with the analytical distributions. More advanced schemes (e.g. implicit schemes) can be found in [21].Methods for the estimation of the parameters of a K-distributed sea clutter can be found in ([2 chapter 9]), in the context of constant false alarm rate detection. In particular, one could compare it to our estimation of α, since α is the shape parameter of the K distribution describing the intensity z_t of the sea clutter (see (11)). Moreover, there is also a different method for estimating $\mathscr{B}$. Indeed, the bandwidth of the intensity is dictated by that of the speckle squared γ_t^2 since the speckle has dynamics much faster than those of the RCS x_t. It appears that the autocorrelation function of

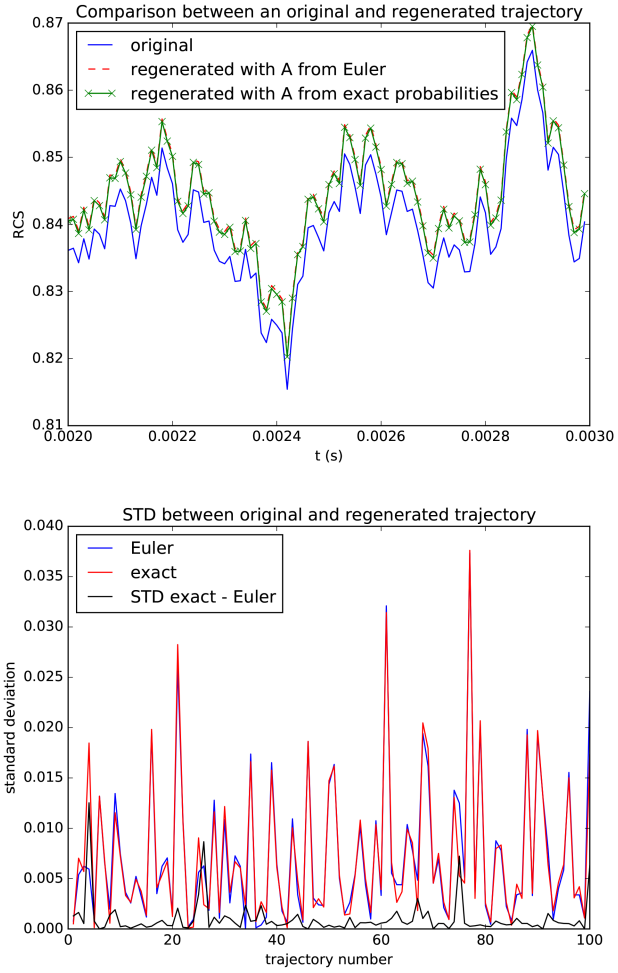


Fig. 5 *Up: example of an original trajectory generated with $\mathscr{A} = 1$ Hz represented along with its regenerated counterparts after estimation of $\mathscr{A}$ using Euler's approximation and the exact transition probabilities. Down: standard deviation of the difference between the original and regenerated trajectory as a function of the trajectory number*

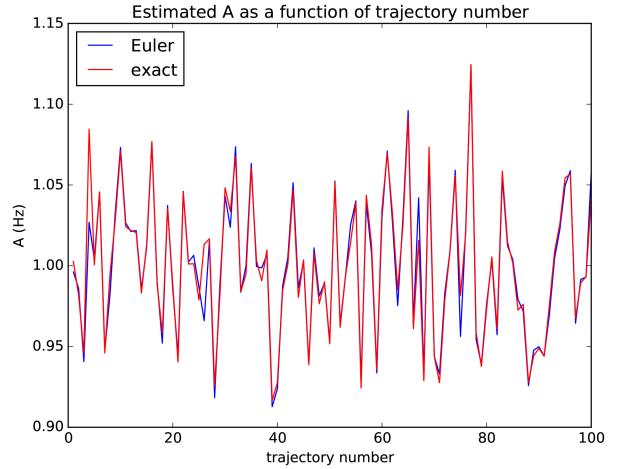


Fig. 6 *Comparison between the estimated $\mathscr{A}$ Euler's approximation and the exact transition probabilities as a function of trajectory number*

γ_t^2 decays as $e^{-\mathscr{B}t}$ (see pp. 63–64 of [3]). As a consequence of the Wiener–Khinchin theorem, the bandwidth of γ_t^2, and so of z_t, varies as $\mathscr{B}$. This remark could be the starting point of another estimation method for $\mathscr{B}$ based on the power spectrum of z_t.

5 Conclusions

The model proposed by Field represents the sea clutter as a solution to SDEs. In this study, we addressed the issue of

estimating its three parameters: $\mathscr{A}$, $\mathscr{B}$ and α, with the support of numerical simulations. $\mathscr{A}$ and $\mathscr{B}$ have been estimated by ML. α has been estimated using the ergodic property of the sea clutter, an estimation both efficient and physically meaningful. We have also proposed a convenient joint estimation of $\mathscr{A}$ and α.

We have compared three methods for the estimation of $\mathscr{A}$ and $\mathscr{B}$. Euler's approximation assumes a constant drift and volatility over small intervals, it leads to an explicit formula for the estimator. Nowman's approximation resolves the drift but assumes constant volatility over small intervals, and leads also to an explicit formula. In contrast, using the exact transition probabilities makes no assumption but does not lead to an explicit formula. It requires numerical minimisation. It was observed in Section 3 that Nowman's method has the same performance as Euler's, so this method was discarded straight away in the rest of the paper.

We have developed a simulation-based method that enables to appreciate and compare the respective performance of estimators of the parameters of a SDE, in connection with the specificities of the application. In our application of measuring the sea-clutter with a satellite system, the specificities were the numerical values of the radiometric accuracy, the time step and the duration of the time series. We proved that even though using the exact transition probabilities gives the best results, Euler's approximation is sufficient provided that the estimator is debiased by indirect inference. From a general perspective, this study emphasises the fact that the statistical estimation issue needs to be addressed by having a thorough knowledge of the operational and physical characteristics of the sensors. With this in mind, we suggest that appropriate studies such as ours could be carried in models based on SDEs to avoid objectively overcomplex estimation methods.

6 Acknowledgments

This work was supported by the DGA, Delegate for Armament of French MoD. We are very grateful to the anonymous reviewers whose comments stimulated us and improved this paper.

7 References

[1] Cumming, I. G., Wong, F. H.: '*Digital processing of synthetic aperture radar data*' (Artech House, Norwood, MA, USA, 2005)

[2] Ward, K., Tough, R., Watts, S., in Stewart, N., Griffiths, M. (Eds.): '*Sea clutter: scattering, the K distribution and radar performance*' (The Institution of Engineering and Technology, Stevenage, UK, 2006), series 20

[3] Field, T. R., in Birman, J., Edwards, S. F., Friend, R., Rees, M., Sherrington, D., Veneziano, G. (Eds.): '*Electromagnetic scattering from random Media*' (Oxford University Press, Oxford, UK, 2009)

[4] Gliklikh, Y. E.: 'Essentials from Stochastic Analysis in Linear Space', in Beiglboeck, W., Chrusciel, P., Eckmann, J.-P., Grosse, H., Kupiainen, A., Loss, M., Löwen, H., Nekrasov, N., Salmhofer, M., Smirnov, S., Takhtajan, L., Yngvason, J. (Eds.): '*Global and stochastic analysis with applications to mathematical physics*' (Springer, Berlin, Germany, 2011), pp. 115–138

[5] Roussel, C. J., Coatanhay, A., Baussard, A.: 'Forward and backward probabilistic inference of the sea clutter', *Waves Random Complex Media*, 2018, doi: 10.1080/17455030.2018.1453956

[6] Nielsen, J. N., Madsen, H., Young, P. C.: 'Parameter estimation in stochastic differential equations: an overview', *Annu. Rev. Control*, 2000, **24**, pp. 83–94

[7] Aït-Sahalia, Y.: 'Maximum likelihood estimation of discretely sampled diffusions: a closed-form approximation approach', *Econometrica*, 2002, **70**, (1), pp. 223–262

[8] Durham, G. B., Gallant, A. R.: 'Numerical techniques for maximum likelihood estimation of continuous-time diffusion processes', *J. Bus. Econ. Stat.*, 2002, **20**, (3), pp. 297–316

[9] Phillips, P. C. B., Yu, J.: 'Maximum likelihood and Gaussian estimation of continuous time models in finance' in Mikoessch, T., Kreiß, J.P., Davis, R., Andersen, T. (Eds.): '*Handbook of financial time series*' (Springer, Berlin, Heidelberg, 2009), pp. 497–530

[10] Ward, K., Baker, C., Watts, S.: 'Maritime surveillance radar. I. Radar scattering from the ocean surface', *IEE Proc. F, Radar Signal Process.*, 1990, **137**, (2), pp. 51–62

[11] Farina, A., Gini, G., Greco, M., *et al.*: 'High resolution sea clutter data: statistical analysis of recorded live data', IEE Proc., Radar Sonar Navig., 1997, **144**, (3), pp. 121–130

[12] Fayard, P., Field, T.: 'Optimal inference of the scattering cross-section through the phase decoherence', *Waves Random Complex Media*, 2008, **18**, (4), pp. 571–584

[13] Oksendal, B.: '*Stochastic differential equations: An Introduction with Applications*' (Springer-Verlag, Berlin, Germany, 2000, 5th edn.)

[14] Nowman, K. B.: 'Gaussian estimation of single-factor continuous time models of the term structure of interest rates', *J. Financ.*, 1997, **52**, (4), pp. 1695–1706

[15] Risken, H., in Haken, H. (Ed.): '*The Fokker–Planck equation: methods of solution and applications*' (Springer, Berlin, Germany, 1989)

[16] Feller, W.: 'Two singular diffusion problems', *Ann. Math.*, 1951, **54**, (1), pp. 173–182

[17] Higham, D. J.: 'An algorithmic introduction to numerical simulation of stochastic differential equations', *Soc. Ind. Appl. Math.*, 2001, **43**, (3), pp. 525–546

[18] Berndt, E. K., Hall, B. H., Hall, R. E., *et al.*: 'Estimation and inference in nonlinear structural models', *Ann. Econ. Soc. Meas.*, 1974, **3**, (4), pp. 653–665

[19] Mai, A. T., Bastin, F., Toulouse, M.: 'On optimization algorithms for maximum likelihood estimation', CIRRELT, Tech. Rep., 2014

[20] Antony, J. W., Schmidt, K., Schwerdt, M., *et al.*: 'Radiometric accuracy and stability of TerraSAR-X and TanDEM-X'. 11th European Conf. on Synthetic Aperture Radar, Hamburg, Germany, 2016

[21] Kloeden, P. E., Platen, E., in Karatzas, I., Yor, M. (Eds.): '*Numerical solution of stochastic differential equations*' (Springer, Berlin, Germany, 1992)

8 Appendix

8.1 Euler–Maruyama and Milstein schemes

Let X_t be a stochastic process solution to the following SDE:

$$\mathrm{d}X_t = \mu(X_t, t)\mathrm{d}t + \sigma(X_t, t)\mathrm{d}W_t, \tag{37}$$

where $(W_t)_t$ is a Brownian motion, also called Wiener process. μ is called the 'drift' and σ is called the 'volatility'. Understanding the algorithm for solving numerically this SDE is a very good way to gain an intuitive understanding of the meaning of a SDE. For precise definitions, e.g. refer to [13]. For what follows, refer [17]. Let $[0, T]$ be a finite time interval and $t_0 = 0 < t_1 < \cdots < t_n$ a partition of $[0, T]$. The Euler–Maruyama method reads:

$$X_{t_i} = X_{t_{i-1}} + \mu(X_{t_{i-1}})(t_i - t_{i-1}) + \sigma(X_{t_{i-1}})(W_{t_i} - W_{t_{i-1}})\,. \tag{38}$$

Equation (38) states that the increment of X between t_{i-1} and t_i is the sum of a term proportional to $t_i - t_{i-1}$ and a term proportional to the increment of the Brownian motion $\Delta W_{t_i} = W_{t_i} - W_{t_{i-1}}$. This increment is a Gaussian random variable with law $\mathcal{N}(0, t_i - t_{i-1})$. Generating a series of n random increments using Python, e.g. leads to one possible trajectory. If many more increment series are generated, we can generate and visualise almost all possible trajectories and evaluate a numerical distribution of X_{t_i} for any i. We also use Mistein's method (see [17]) when necessary. Milstein scheme reads

$$\begin{aligned} X_{t_i} = X_{t_{i-1}} &+ \mu(X_{t_{i-1}})(t_i - t_{i-1}) + \sigma(X_{t_{i-1}})(W_{t_i} - W_{t_{i-1}}) \\ &+ \frac{1}{2}\sigma(X_{t_{i-1}})\sigma'(X_{t_{i-1}})((W_{t_i} - W_{t_{i-1}})^2 - (t_i - t_{i-1})) \end{aligned} \tag{39}$$

and it has a higher order of convergence.