

Estimation of marginal parameters of SUP-OU processes with long range dependence

Emanuele Taufer^{1*}

¹Department of Economics and Management Sciences, University of Trento

*Email: emanuele.taufer@unitn.it

Abstract

Superpositions of Ornstein Uhlenbeck processes provide convenient ways to build stationary processes with given marginal distributions and long range dependence. After reviewing some of the basic features, we present several examples of processes with non Gaussian marginal distributions. Estimation of the parameters of the marginal distribution is undertaken by means of a characteristic function technique. We provide the relevant asymptotic theory as well as results of simulations and real data applications.

Keywords: Characteristic function estimation; long range dependence; marginal distribution; Ornstein Uhlenbeck process; superposition.

1. Introduction

The modelling of stochastic processes with given marginal distributions and tractable dependence structures can be a point of considerable interest in fields of applications such as finance and econometrics where, for example, the identification of models for assets prices which are consistent with some stylized features such as heavy tails, asymmetry and long range dependence (LRD) is quite important as a first step to pricing derivatives and general risk management.

To this end, the use of stationary processes of the Ornstein Uhlenbeck (OU) type and their superpositions (Sup OU) has been thoroughly studied by Barndorff-Nielsen (1998, 2001), Barndorff-Nielsen, Jensen and Sorensen (1998), Barndorff-Nielsen and Shephard (2001, 2002); a good reference is also Schoutens (2003). Earlier related works which exploits the use of superpositions as a way of modeling long range dependence are those of Granger (1980), Cox (1991), Ding and Granger (1996), Comte and Renault (1998).

In particular, the use of self decomposable random variables allows one to obtain Sup OU processes with specified marginal distributions and, at the same time, given correlation structures; furthermore, specific constructions allow one to obtain LRD. Most notable examples include Sup OU processes with weak or long range dependence and marginal distributions such as the Normal Inverse Gaussian and the Inverse Gaussian (Barndorff-Nielsen, 1998), the Variance Gamma (Seneta, 2004), the Meixner (Schoutens and Teugel, 1998), the Normal and the Gamma.

Leonenko and Taufer (2005) have provided an invariance principle for Sup OU processes with LRD; the limiting process obtained, a fractional Brownian motion, connects this approach to the mainstream ones of the statistical literature which are based on non-linear transformations of Gaussian LRD processes and linear processes with LRD. For other related papers see Anh et al. (2010), Leonenko

et al. (2013), Leonenko and Taufer (2013).

In this paper we will concentrate on estimation of the marginal parameters in Sup OU processes. The main focus will be on Sup OU processes with LRD. However, the technique adopted, based on the empirical characteristic function (ECF) will be seen to work for finite Sup OU processes as well. Estimation of the dependence structure can be tackled as discussed in Leonenko and Taufer (2005).

Estimation based on the ECF is well established and a good starting point is the paper of Feuerverger and McDunnough (1981) who show that arbitrarily high levels of efficiency can be obtained by such methods in the i.i.d. case; furthermore, they discuss the extension to dependent observations, other references that are relevant to the subsequent development are those of Madan and Seneta (1987), Knight and Satchell (1997), Knight and Yu (2002), Taufer and Leonenko (2009) and Taufer et al. (2011) which discuss ECF estimation for OU processes and stochastic volatility models. Other contributions on estimation of OU processes are those of Woerner (2004), Baran, Pap, and van Zuijlen (2003), Gloter (2001), Florens-Landais and Pham (1999), Pap and van Zuijlen (1996).

Barndorff-Nielsen (1998) has shown how to implement likelihood analysis for discretely observed OU processes. However, when we come to Sup OU processes it is not at all straightforward to apply the methods discussed in the references above. Indeed, likelihood analysis of Sup OU processes does not seem feasible at all. As we will see, estimation based on the empirical characteristic function fits quite naturally into the framework of Sup OU processes. Moreover, the adaptability of ECF estimation method to either i.i.d. or dependent case is a point of considerable importance in financial applications where usually estimation of the marginal parameters is undertaken on the assumption of independence of returns. This may not be often justifiable in practice where GARCH effects are the rule more than the exception. This fact was noted also by Seneta(2004) which

argues that moment based estimation may be less sensitive with respect to maximum likelihood to departures from the independence hypothesis.

Last, but not least, several models that are often used in practice as the Normal Inverse Gaussian or the Variance Gamma do require special care when using maximum likelihood estimation since the densities involve modified Bessel functions, as we will see, ECF estimation will avoid these problems and numerically fast algorithms are available.

2. Background

The present section reviews some known results; for further details the reader is referred to Wolfe (1982), Barndorff Nielsen et al. (1998) and Barndorff Nielsen (1998, 2001), Leonenko and Taufer (2005) in which stationary processes of Ornstein Uhlenbeck (OU) and their superpositions (Sup OU) with weak or long range dependence and self decomposable marginal distributions are constructed.

Recall that a random variable X with characteristic function (CF) $\phi(\zeta)$ is self decomposable if, for all $c \in (0, 1)$, there exists a CF $\phi_c(\zeta)$ such that $\phi(\zeta) = \phi(c\zeta)\phi_c(\zeta)$.

An OU process $\{X(t), t \geq 0\}$ satisfies the stochastic differential equation

$$dX(t) = -\lambda X(t)dt + d\dot{Z}(t), \tag{2.1}$$

where $\lambda > 0$ and $\dot{Z}(t)$ is a homogeneous Lévy process for which $E[\log(1 + |\dot{Z}(1)|)] < \infty$; it is commonly referred to as the background driving Lévy process (BDLP). Equation (2.1) has the solution

$$X(t) = e^{-\lambda t} X(0) + \int_0^t e^{-\lambda(t-s)} d\dot{Z}(s). \tag{2.2}$$

Let $\phi(\zeta)$ and $\kappa(\zeta) = \log \phi(\zeta)$ be respectively the CF and the cumulant function of X . If $X(t)$ is to be stationary, then it must be that $\phi(\zeta) = \phi(\zeta e^{-\lambda t})\phi_t(\zeta)$ for all $t \geq 0$, where $\phi_t(\zeta)$ denotes the CF of the second term in the r.h.s. of (2.2).

If $\zeta \kappa'(\zeta)$ can be defined at 0 by continuity, then (Barndorff-Nielsen et al. 1998, Lemma 3.1) $\exp\{\zeta \kappa'(\zeta)\}$ is an infinitely divisible CF. It follows that the choice $\phi(\zeta) = \lambda \zeta \kappa'(\zeta)$ where $\exp\{\phi(\zeta)\}$ is the CF of $\dot{Z}(1)$ implies that

$$\phi_t(\zeta) = \exp\left\{\int_{\zeta e^{-\lambda t}}^{\zeta} \kappa'(w)dw\right\} \tag{2.3}$$

from which, by continuity, as $\phi(\zeta) = \lim_{t \rightarrow \infty} \phi_t(\zeta)$, it follows that $\phi(\zeta)$ is the CF of the marginal distribution of the OU process (2.2).

Remark 1. The stationary process $X(t), t \geq 0$ can be extended to a stationary process on the whole real line as indicated in Barndorff-Nielsen (1995) by introducing an independent copy of the process $\dot{Z}(t)$ but modifying it to be cadlag.

As noted by Wolfe (1982), for self decomposable distributions, a discrete AR(1) process

$$X_{n+1} = cX_n + \varepsilon_{n+1}, \quad n \geq 1 \tag{2.4}$$

can be embedded into the OU process (2.2) if we set $c = e^{-\lambda}$ and $\varepsilon_n \stackrel{D}{=} Z(1)$ where,

$$Z(1) = e^{-\lambda} \int_0^1 e^{\lambda s} d\dot{Z}(s). \tag{2.5}$$

Note in fact that, from (2.3), we can make the estimate

$$E(e^{i\zeta Z(1)}) = \exp\left\{\int_{\zeta e^{-\lambda}}^{\zeta} \kappa'(w)dw\right\} = \exp\{\kappa(\zeta) - \kappa(\zeta e^{-\lambda})\}. \tag{2.6}$$

and if ε_n has CF corresponding to (2.6) then X_n has CF $\phi(\zeta)$. The processes X_n and $X(t)$, sampled at $n, t = 0, 1 \dots$ will have the same marginal distribution. Given this paper will concentrate on estimation of discretely sampled OU or AR(1) processes we will use the following common definition.

Definition 1. $X_n, n \geq 0$ will be defined an OU process with CF $\phi(\zeta)$ if it satisfies (2.2) with $t = 0, 1 \dots$ or (2.4) with $c = e^{-\lambda}$.

The sum of independent OU processes can be exploited in order to obtain processes with given marginal distribution and more elaborated correlation structures. Suppose we have a sequence of independent OU processes of the form

$$dX^{(k)}(t) = -\lambda_k X^{(k)}(t) dt + d\dot{Z}^{(k)}(t), \quad k = 1, 2, \dots \tag{2.7}$$

From the discussion above note that the choice $\phi_k(\zeta) = \delta_k \lambda_k \zeta \kappa'(\zeta)$, where $\exp\{\phi_k(\zeta)\}$ is the CF of $\dot{Z}^{(k)}(1)$, implies that the CF of $X^{(k)}(t)$ is $\phi(\zeta)^{\delta_k}$. Then, for an integer M

$$X(t) = \sum_{k=1}^M X^{(k)}(t) \tag{2.8}$$

has marginal distribution with CF $\phi(\zeta)^\delta$, where $\delta = \sum_{k=1}^M \delta_k$. Letting $\sigma^2 = \text{Var}(X(t))$,

$$\text{Cov}(X(t), X(t + \tau)) = \sigma^2 \sum_{k=1}^M \delta_k e^{-|\tau|\lambda_k} \quad \tau \in \mathbb{R} \tag{2.9}$$

From Barndorff-Nielsen (2001) the superposition stretched to the case $M = \infty$ provides well defined stationary process in the sense of mean square or almost sure convergence. In this latter case, for a constant c and $H \in (0, 1)$, the choice $\delta_k = ck^{-1-2(1-H)}$, and $\lambda_k = \lambda/k$, implies that the correlation function (2.8) satisfies

$$L(u)u^{-2(1-H)} \tag{2.10}$$

where $L(u)$ is slowly varying at infinity. Thus, if $H \in (1/2, 1)$, $X(t)$ exhibits LRD. Note that the same features will be present in the superposition of discrete AR(1) processes. Without loss of generality we may consider $\delta = 1$, and hence we will use the following definitions:

Definition 2. $X(t)$ satisfying (2.8) with finite M and marginal distribution with CF $\phi(\zeta)$, having finite variance σ^2 , will be defined Sup OU process. The notation $X_n, n = 0, 1 \dots$ will be used when $X(t)$ is sampled at instants $t = 0, 1 \dots$.

Definition 3. $X(t)$ satisfying (2.8) with infinite M , $\delta_k = ck^{-1-2(1-H)}$, and $\lambda_k = \lambda/k$, c a constant and $H \in (0.5, 1)$, having marginal distribution with CF $\phi(\zeta)$ and finite variance σ^2 , will be defined LRD Sup OU process. The notation $X_n, n = 0, 1 \dots$ will be used when $X(t)$ is sampled at instants $t = 0, 1 \dots$.

Leonenko and Taufer (2005) have provided an invariance principle for sums and partial sums of Sup OU and LRD Sup OU processes. For the purposes of the methods studied in this paper it will suffice to consider the following

Proposition 1. Let $S_N = \sum_{n=1}^N (X_n - EX_n)$ then we have asymptotic normality of:

- $N^{-1/2} S_N$ if $X_n, n = 1 \dots N$, is a Sup OU process;
- $N^{-H} L(N)^{-1/2} S_N$ if $X_n, n = 1 \dots N$, is an LRD Sup OU process and $L(\cdot)$ a slowly varying function at infinity.

3. Characteristic function estimation

We will consider a situation where we have a sample $X_n, n = 1, \dots, N$ of a (LRD) Sup OU process with marginal distribution F_θ , where $\theta \in \Theta \subset \mathbb{R}^p, p \in \mathbb{N}$, is a vector of parameters. Our main aim will be to estimate the parameter vector θ .

Note that we are concerned with estimation of the parameters of the marginal distribution of the process and not about the long memory parameter H ; this may be tackled, for example, as discussed in Leonenko and Taufer (2005).

In order to stress the dependence on a parameter vector θ we will denote by $\phi_\theta(u)$ the CF of the random variable with distribution F_θ .

3.1. A simple procedure

Let $X_n, n = 1, \dots, N$ be a (LRD) Sup OU process with marginal distribution having CF $\phi_\theta(u)$ where θ has unknown true value θ_0 . Let $0 < u_1 < \dots < u_q$ be a fixed grid; define the empirical counterpart of $\phi_\theta(u)$ as $\hat{\phi}_N(u)$. To estimate the parameters of the marginal distribution, choose θ to minimize the quadratic form

$$\hat{\theta}_N = \arg \min_{\theta} Q_N(\theta), \quad Q_N(\theta) = \sum_{j=1}^q |\hat{\phi}_N(u_j) - \phi_\theta(u_j)|^2. \quad (3.11)$$

We will implicitly assume that the parametrization used brings to an identifiable problem, i.e. $\phi_{\theta'}(u) \neq \phi_\theta(u)$ if $\theta' \neq \theta$. Given this, to prove consistency and asymptotic normality of the proposed procedure, we need some standard regularity conditions which we list here.

A1. Let the parameter space Θ be a compact set and $\theta_0 \in \text{Int}(\Theta)$ and assume $\phi_\theta(u)$ is continuously differentiable in θ .

A2. The matrix $B(\theta) = \sum_{j=1}^q \frac{\partial \phi_\theta(u_j)}{\partial \theta} \frac{\partial \phi_\theta(u_j)}{\partial \theta'}$ is of full rank.

Theorem 1. Let $X_n, n = 1 \dots N$, be a Sup OU process or a LRD Sup OU process. Then, under A1, $\hat{\theta}_N \rightarrow \theta$ a.s..

Theorem 2. Let $A_N(\theta) = \frac{1}{2} \frac{\partial Q_N(\theta)}{\partial \theta}$. Then, under A1 and A2,

- if $X_n, n = 1 \dots N$, is a Sup OU process,

$$N^{1/2} (\hat{\theta} - \theta_0) \xrightarrow{D} \mathcal{N} \left(0, B(\theta_0)^{-1} A(\theta_0) B(\theta_0)^{-1} \right). \quad (3.12)$$

with $A(\theta) = \lim_{N \rightarrow \infty} N^{-1} E[A_N(\theta)A_N(\theta)']$

- if $X_n, n = 1 \dots N$, is an LRD Sup OU process,

$$N^{1-H} L(N)^{-1/2} (\hat{\theta} - \theta_0) \xrightarrow{D} \mathcal{N} \left(0, B(\theta_0)^{-1} A^*(\theta_0) B(\theta_0)^{-1} \right). \quad (3.13)$$

with $A^*(\theta) = \lim_{N \rightarrow \infty} N^{-2H} L(N)^{-1} E[A_N(\theta)A_N(\theta)']$

Remark 2. It is clear that the estimator proposed will work also in a i.i.d. setting providing a consistent and asymptotically normal procedure. For the asymptotic variance in such a case the reader is referred to Feuerverger and McDunnough (1981).

Table 1 gives some distributions which are self decomposable and can be exploited in order to construct (LRD) Sup OU processes with given marginal distribution by superposition of independent OU processes.

Together with common distributions such as the Normal and the Gamma, we find some less known examples such as the Simmetric Gamma, for which the reader is referred to Dufresne (1997), Kotz et Al. (2001, p. 179) and Steutel and van Harn (2004, p. 504), the Meixner, for which one can consult Schoutens and Teugel (1998) and Grigelionis (1999), the Normal Inverse Gaussian and the Inverse Gaussian which are well known in applications to financial data, see Barndorff-Nielsen (1997, 1998), Rydberg (1997), Barndorff-Nielsen and Shephard (2001); and the Variance Gamma distribution which provides a suitable model for financia data as well, as discussed, for example, in Madan and Seneta (1990) and Seneta (2004), a general reference text on self decomposable distributions and financial applications is Schoutens (2003).

Table 1: Distributions which may be used to construct (LRD) Sup OU processes

Name and Reference	log CF
Normal $N(\mu_k, \delta_k)$ $x \in \mathbb{R}$	$\mu i \zeta - \delta \zeta^2 / 2$
Norm. Inv. Gaussian $NIG(\alpha, \beta, \delta_k, \mu_k)$ $x \in \mathbb{R}$	$i \mu \zeta - \delta (\sqrt{\alpha^2 - (\beta + i \zeta)^2} - \sqrt{\alpha^2 - \beta^2})$
Inverse Gaussian $IG(\alpha_k, \beta)$ $x > 0$	$-\alpha (\sqrt{-2i \zeta + \beta^2} - \beta)$
Gamma $\Gamma(\alpha_k, \beta)$ $x > 0$	$\log \left(1 - \frac{i \zeta}{\beta} \right)^{-\alpha}$
Variance Gamma $VG(\alpha, \beta, \delta_k, \mu_k)$ $x \in \mathbb{R}$	$\mu i \zeta + 2 \delta \log \left(\frac{\sqrt{\alpha^2 - \beta^2}}{\sqrt{\alpha^2 - (\beta + i \zeta)^2}} \right)$
Sym. Gamma $S\Gamma(\alpha_k, \beta)$ $x \in \mathbb{R}$	$\log \left(1 + \frac{\zeta^2}{\beta^2} \right)^{-\alpha}$
Meixner $M(\alpha, \beta, \delta_k, \mu_k)$ $x \in \mathbb{R}$	$i \mu \zeta + 2 \delta \log \left(\frac{\cos(\beta/2)}{\cosh((\alpha \zeta - i \beta)/2)} \right)$

3.2. Use of weighting functions and continuous procedures

The procedure (3.11) lends itself to further manipulation by considering weighting functions depending on the parameter, i.e. $w_\theta(u)$ or, more generally, in order to avoid the choice of a grid of points $u_j, j = 1, \dots, q$ by considering a continuous version

$$\int |\hat{\phi}_N(u) - \phi_\theta(u)|^2 dG(u) \quad (3.14)$$

where $G(u)$ is a suitable weighting function. It is clear that if $G(u)$ is a step function the we are actually considering again the discrete case.

For a discussion of the choice of optimal weights, see Feuerverger and McDunnough (1981) which show that, in the i.i.d. case, using a weight given by the Fourier transform of the score is asymptotically equivalent to maximum likelihood. This reasonig can be generalized to dependent observations, however it is unfeasible to compute the given quantities in the case of LRD Sup OU processes. Other weighting schemes may be obtained by constructing linear or quadratic forms involving weighting matrices W_θ connected to the autocovariance matrix of $\phi_N(u)$. Again, the evaluation of the the asymptotic autocovariance matrix of $\phi_N(u)$ may be problematic for LRD Sup OU processes, as we will see in the proof of Lemma 2. In the next sub section we will provide a delete- d jackknife estimator of the variance of $\hat{\theta}$ which may be adapted for this purpose.

If we consider avoiding the choice of a grid and use a continuous version of (3.11), the choice of a proper weighting function $G(u)$ is important in order to have a finite integral and for computability reasons. Knight and Yu (2002) discuss the use of and exponential weight, i.e. $dG(u) = \exp\{-u^2\} du$, Epps (2005), in the context of goodness of fit testing, suggests using $dG(u) = (|\phi_\theta(u)|^2 / \int |\phi_\theta(v)|^2 dv) du$. Both choices have the effect of damping out the persistent oscillations of $|\hat{\phi}_N(u) - \phi_\theta(u)|$, as $u \rightarrow \infty$ assuring the finiteness of (3.14) but may not be optimal for efficiency considerations.

Regarding asymptotic properties of these generalizations, it is clear that Theorem 1 and Theorem 2 can be readily adapted when consid-

Table 2: Mean and MSE of estimates based on simulated data of independent observations, 1000 iterations. Moment estimation (ME) and CF estimation (CFE).

n	ME		CFE		
	Mean	MSE	Mean	MSE	
100	$\alpha = 2$	3.7316	7.9624	1.8271	0.5359
	$\beta = 1$	2.2773	6.2362	0.7047	0.2870
	$\delta = 1$	1.2690	0.2710	1.0066	0.1005
	$\mu = 1$	0.6970	0.3230	1.1404	0.0641
500	$\alpha = 2$	3.1783	5.0184	2.0459	0.3105
	$\beta = 1$	2.0072	3.9547	1.0269	0.1655
	$\delta = 1$	1.1135	0.0829	1.0009	0.0292
	$\mu = 1$	0.7357	0.2097	1.0027	0.0250
1000	$\alpha = 2$	2.7876	2.4466	2.1017	0.2314
	$\beta = 1$	1.6595	1.7960	1.0744	0.1345
	$\delta = 1$	1.0927	0.0521	1.0156	0.0164
	$\mu = 1$	0.8028	0.1344	0.9781	0.0183

ering the minimization of a weighted form over a finite grid of points with no change in the starting conditions except the one that the weight exists over the points considered. If we consider minimizing (3.14), then, Theorem 1 still holds under the conditions discussed for the discrete case, see for example Knight and Yu (2002), for applying the methods of Theorem 2 we need to impose that the function $K_j(X_n, \theta) = K(X_n, \theta, u_j)$, defined in the proof of Theorem 2, is such that

$$\int K(X_n, \theta, u) dG(u) \tag{3.15}$$

is a bounded measurable function of X_n with bounded first derivative for all θ . We note that a similar condition is used in Knight and Yu (2002).

4. An example with simulated data, the NIG case

To have an idea of the performance of the estimators proposed for some important classes of distributions, we consider a small simulation study which has the double aim of addressing practical issues and to show what happens when we pass from a situation of independence to one of dependence.

We generated samples of various sizes from a NIG(2,1,1,1) distribution either in a i.i.d. setting as well as for an OU process. Simulation of OU processes can be done as described in Taufer and Leonenko (2009b).

Table 1 gives the mean and the MSE of moment estimators (ME) and ECF estimators (ECFE) in the i.i.d. case while Table 2 considers the case where the data represent an OU process with $\lambda = 1$.

ME are to be found, for example in Karlis (2002) and we note that, depending on sample realization, they may not exist. The 1000 iterations of Table 2 have been obtained for cases where the ME exist, we discarded those replications for which they did not; this phenomenon was, intuitively, more present in small sample sizes: for $n = 100$ we had to discard more than 400 samples before reaching the 1000 replicates.

Placing the attention on the results, we see that the bias of ME for α and β which are connected to the third and fourth moment of the distribution may be substantial and remains quite high even in large sample sizes. The problem is less serious for the other two parameters, μ and δ which are connected to location and variability. ECFE perform much better either in terms of bias than MSE.

The situation changes quite a bit in the case of an OU process. Here ME are highly unreliable even for sample sizes as high as 1000 and 3000 and there were several instance in which they did not exist. In Table 3 such values are not reported and note that bias and MSE

Table 3: Mean and MSE of estimates based on simulated data from an OU process with $\lambda = 1$, 100 iterations. Moment estimation (ME) and CF estimation (CFE).

n	ME		CFE		
	Mean	MSE	Mean	MSE	
1000	$\alpha = 2$	-	-	2.7452	0.8716
	$\beta = 1$	-	-	1.6776	0.6982
	$\delta = 1$	-	-	1.0410	0.0186
	$\mu = 1$	-	-	0.8079	0.0547
3000	$\alpha = 2$	-	-	2.4831	0.3622
	$\beta = 1$	-	-	1.4074	0.2365
	$\delta = 1$	-	-	1.0309	0.0081
	$\mu = 1$	-	-	0.8764	0.0226
5000	$\alpha = 2$	4.9332	8.6126	2.2680	0.1514
	$\beta = 1$	3.3924	5.7440	1.1745	0.0750
	$\delta = 1$	1.3252	0.1085	1.0383	0.0064
	$\mu = 1$	0.3241	0.4572	0.9463	0.0086

are quite high for $n = 5000$ as well. These sample sizes may not be considered unnaturally high given the quantity of data that may be available for financial applications. As we see, ECFE suffer much less from this situation. Although bias and MSE are worse w.r.t. the i.i.d. case, we see that the situation is much better. The ECFE have been obtained for $q = 80$ and a fixed step $\tau = u_{i+1} - u_i, i = 1 \dots q$ of size 0.05. The choice of the step size maybe quite important in order to minimize effectively the distance of the ECFE and the fitted one. Feuerverger and McDunnough (1981) discuss optimizing procedures in order to obtain, for given q , either a optimal fixed step τ of optimal points u_1, \dots, u_q . Knight and Yu (2002) avoid this problem by considering continuous CF estimation. Given the expressions for the variance of the ECFE in the case of Sup OU or LRD Sup OU processes, the minimizing procedures to compute a optimal step τ may be quite cumbersome. We run several simulations for different parameter values and found very effective the heuristic approach of plotting the ECF and choose the step τ which samples the whole range of the ECF where is present most of its variability. As far as the sample size q , generally, a larger value seems to increase precision, although it seems that little is gained for $q > 80$. These results are in line with those of Feuerverger and McDunnough (1981).

We run other simulations for the $IG(\alpha, \beta)$ and the $VG(\alpha, \beta, \delta, \mu)$ obtaining similar results. The estimation of α and β in the VG case have the same behaviour as the NIG case. For the VG case, our procedure works well, while we must note that Madan and Seneta (1987) reported bad behaviour of the CF estimators; note however that in their case they were using only the real part of the CF.

5. A real data example

As an application to real data we consider the daily returns of the Mediaset share on the Italian Stock market, from 1st Jan. 1996, to 25 Dec. 2000, for a total of 1300 observations. Although the returns on shares are often considered to be independent, this data set was thoroughly investigated by De Bona (2005) which applies several procedures to detect the presence of long memory effects in the Italian Stock market. For this data set, all tests applied are significant on the presence of long memory effects. Hence one may try to fit a LRD Sup OU model with NIG marginal distribution.

Given our aim is not a full analysis of this data set, here we concentrate on the marginal parameters only. Figure 1 shows the empirical and the fitted ch.f. for the data; the ECF estimation procedure applied with $q = 80$ and $\tau = 0.35$ obtained a $NIG(12.60, 2.52, 0.08, -0.0004)$: as we see from Figure 2, the PP-plot of the data shows there is nearly a perfect fit of the marginal distribution.

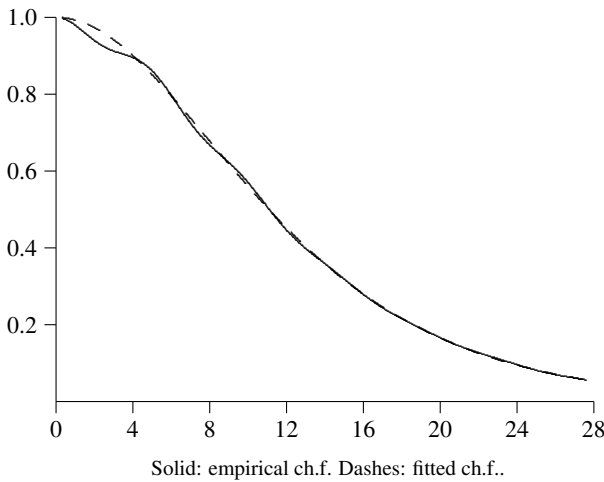


Figure 1: Empirical and fitted ch.f. of the Mediaset data

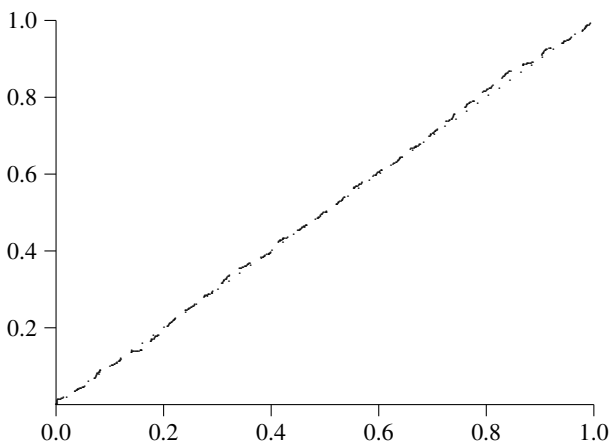


Figure 2: PP plot of Mediaset Data fitted by the NIG distribution

6. Conclusions

We have discussed the use of ECF estimation in the context of LRD: we have seen that this approach fits quite naturally to LRD Sup OU processes for which other methods seem hard to adapt.

The method proposed has the advantage that the estimation procedure needs not to be modified for Sup OU processes with weak dependence or even for the i.i.d. case; moreover, fast numerical algorithms are available.

This feature may be quite important in practical applications. For example the independence of returns of financial assets is often assumed, although this may be rarely justified since it is not uncommon, to observe (strongly) correlated squared or absolute returns. Seneta (2004) noted the sensitivity of maximum likelihood estimates to these effects, suggesting to use moment estimators which may be more robust to departures from the independence of returns hypothesis. In this sense, the use of CF estimation may be a valid alternative which may be used safely in presence of different dependence structures.

7. Proofs

We will state a couple of preliminary results which we will put in the form of lemmas.

Lemma 1. Let $h(\cdot)$ be a bounded function with bounded first derivative, and let $\chi(X_1, \dots, X_p)$ denote the cumulant of X_1, \dots, X_p . There exists a universal constant C such that

$$\chi(h(X_1), \dots, h(X_p)) \leq C|\chi(X_1, \dots, X_p)| \tag{7.16}$$

Proof. We can expand $h(x) = h(0) + xh'(\xi)$, $|\xi| \leq |x|$, given h and h' are bounded, by the multilinearity property of cumulants we can take the estimate

$$\begin{aligned} \chi(h(X_1), \dots, h(X_p)) &= \chi(h(0) + X_1h'(\xi_1), \dots, h(0) + X_ph'(\xi_p)) \\ &= [h'(\xi_1) \dots h'(\xi_p)]\chi(X_1, \dots, X_p) \\ &\leq C|\chi(X_1, \dots, X_p)| \end{aligned} \tag{7.17}$$

□

Lemma 2. Let X_n be a LRD Sup OU process with $\delta_k = ck^{-1-2(1-H)}$, c a constant such that $\sum_k \delta_k = 1$, $\lambda_k = \lambda/k$, $H \in (1/2, 1)$. Define $\phi(u) = Ee^{iuX_n}$ and suppose $\phi'(\xi)$ can be defined for $\xi \in [\min(0, u, v), \max(0, u, v)]$. Then, as $|n - n'| \rightarrow \infty$,

$$\begin{aligned} Ee^{iuX_n + ivX_{n'}} &= \begin{cases} \phi(u)\phi(v)\exp\{uL_1(|n - n'|)|n - n'|^{-2(1-H)}\} & n > n', \\ \phi(u)\phi(v)\exp\{vL_2(|n - n'|)|n - n'|^{-2(1-H)}\} & n < n' \end{cases} \end{aligned} \tag{7.18}$$

where L_1 and L_2 are slowly varying functions at infinity.

Proof. Clearly, since $X_n = \sum_{k=1}^{\infty} X_n^{(k)}$ is a superposition of independent OU processes, we have

$$E\exp\{iuX_n + ivX_{n'}\} = \prod_{k=1}^{\infty} E\exp\{iuX_n^{(k)} + ivX_{n'}^{(k)}\} \tag{7.19}$$

Let us analyze in more detail the terms in the infinite product, suppose $n > n'$, given $X_n^{(k)}$ is an OU (AR1) process, it holds

$$X_n^{(k)} = e^{-\lambda(n-n')/k}X_{n'}^{(k)} + \sum_{j=0}^{n-n'-1} e^{-\lambda j/k}\epsilon_{n-j}^{(k)} \tag{7.20}$$

and hence, by independence, we can take the estimate

$$\begin{aligned} & \text{Eexp}\{iuX_n^{(k)} + ivX_{n'}^{(k)}\} \\ &= \text{Eexp}\{i(v + ue^{-\lambda(n-n')/k})X_{n'}^{(k)}\} \prod_{j=0}^{n-n'-1} \text{Eexp}\{ive^{-\lambda j/k} \varepsilon_{n-j}^{(k)}\} \end{aligned} \tag{7.21}$$

Recall from Remark 1 that $\varepsilon_n^{(k)} \stackrel{D}{=} Z^{(k)}(1)$ and hence, formal calculation from (2.6) with $\text{E}e^{iuX_n^{(k)}} = \phi(u)^{\delta_k} = \exp\{\delta_k \kappa(u)\}$ obtains that the characteristic function of $\varepsilon_{n-j}^{(k)}$, evaluated at u is $\exp\{\delta_k[\kappa(u) - \kappa(ue^{-\lambda j/k})]\}$, and then

$$\begin{aligned} & \prod_{j=0}^{n-n'-1} \text{Eexp}\{iue^{-\lambda j/k} \varepsilon_{n-j}^{(k)}\} \\ &= \prod_{j=0}^{n-n'-1} \exp\{\delta_k[\kappa(ue^{-\lambda j/k}) - \kappa(ue^{-\lambda(j+1)/k})]\} \\ &= \exp\left\{\sum_{j=0}^{n-n'-1} \delta_k[\kappa(ue^{-\lambda j/k}) - \kappa(ue^{-\lambda(j+1)/k})]\right\} \\ &= \exp\{\delta_k[\kappa(u) - \kappa(ue^{-\lambda(n-n')/k})]\} \end{aligned} \tag{7.22}$$

Putting together the above expressions we have the estimate

$$\begin{aligned} & \text{Eexp}\{iuX_n^{(k)} + ivX_{n'}^{(k)}\} \\ &= \exp\{\delta_k[\kappa(v + ue^{-\lambda(n-n')/k}) + \kappa(u) - \kappa(ue^{-\lambda(n-n')/k})]\} \end{aligned} \tag{7.23}$$

Now, by the mean value theorem, we have that, for $|\xi_1^{(k)} - v| \leq |ue^{-\lambda(n-n')/k}|$, $\kappa(v + ue^{-\lambda(n-n')/k}) = \kappa(v) + ue^{-\lambda(n-n')/k} \kappa'(\xi_1^{(k)})$ and, for $|\xi_2^{(k)}| \leq |ue^{-\lambda(n-n')/k}|$, $\kappa(ue^{-\lambda(n-n')/k}) = \kappa(0) + ue^{-\lambda(n-n')/k} \kappa'(\xi_2^{(k)})$. Since $\kappa(0) = 0$,

$$\begin{aligned} & \text{Eexp}\{iuX_n^{(k)} + ivX_{n'}^{(k)}\} \\ &= \phi(u)\phi(v) \exp\{\delta_k[ue^{-\lambda(n-n')/k}[\kappa'(\xi_1^{(k)}) - \kappa'(\xi_2^{(k)})]]\} \end{aligned} \tag{7.24}$$

Going back to our original goal we have that $\text{E}e^{iuX_n + ivX_{n'}}$ is given by

$$\phi(u)\phi(v) \exp\left\{u \sum_{k=1}^{\infty} \delta_k e^{-\lambda(n-n')/k} [\kappa'(\xi_1^{(k)}) - \kappa'(\xi_2^{(k)})]\right\} \tag{7.25}$$

We now need to evaluate further the infinite sum appearing in the exponent. To this end note that $\kappa'(\xi_1^{(k)}) - \kappa'(\xi_2^{(k)}) \rightarrow \kappa'(\xi_1) - \kappa'(\xi_2)$, a constant, as $k \rightarrow \infty$, where $|\xi_1 - v| \leq |u|$ and $|\xi_2| \leq |u|$. The terms in the sequence will be decreasing after a certain point and then, for $|n - n'| \rightarrow \infty$, we can approximate the sum by the integral

$$\int_1^{\infty} \frac{c}{y^{1+2(1-H)}} e^{-\lambda|n-n'|/y} [\kappa'(\xi_1^{(y)}) - \kappa'(\xi_2^{(y)})] dy \tag{7.26}$$

apply the transformation $\lambda|n - n'|/y = w$ to obtain the following estimate

$$c(\lambda|n - n'|)^{-2(1-H)} \int_0^{\lambda|n-n'|} w^{2(1-H)-1} e^{-w} L(w) dw \tag{7.27}$$

where $L(w) = [\kappa'(\xi_1^{(w)}) - \kappa'(\xi_2^{(w)})]$ with $|\xi_1^{(w)} - v| \leq |ue^{-w}|$ and $|\xi_2^{(w)}| \leq |ue^{-w}|$, note that $\lim_{w \rightarrow \infty} L(cw)/L(w) = 1$, $c > 0$ and hence L is slowly varying at infinity. Hence, for $|n - n'| \rightarrow \infty$ the above term can be evaluated as

$$\begin{aligned} & c(\lambda|n - n'|)^{-2(1-H)} L(\lambda|n - n'|) \int_0^{\lambda|n-n'|} w^{2(1-H)-1} e^{-w} dw \\ &= \frac{L_1(|n - n'|)}{|n - n'|^{2(1-H)}} \end{aligned} \tag{7.28}$$

We note that $\lim_{w \rightarrow \infty} L(w) = \kappa'(v) - \kappa'(0)$ and the integral in the expression above tends to $\Gamma(2(1 - H))$ integral. The case $n' > n$ can be treated in a similar way. \square

Proof of Theorem 1. Let $Q_0(\theta) = \sum_{j=1}^q |\phi_{\theta_0}(u_j) - \phi_{\theta}(u_j)|^2$, then, as in Knight and Yu (2002, Theorem 2.1) we can take the estimate

$$|Q_N(\theta) - Q_0(\theta)| \leq 4 \sum_{j=1}^q |\phi_N(u_j) - \phi_{\theta_0}(u_j)| \tag{7.29}$$

Since $\phi_N(u) \rightarrow \phi_{\theta}(u)$ a.s. by the ergodic property, consistency of the estimator follows by standard arguments under the assumptions of the theorem. \square

Proof of Theorem 2. The first step of the proof will be to show that $\frac{\partial Q_N(\theta)}{\partial \theta}$ properly normalized, converges to a Normal random variable. To this end note that

$$\frac{\partial Q_N(\theta)}{\partial \theta} = -\frac{2}{N} \sum_{n=1}^N \sum_{j=1}^q K_j(X_n, \theta) \tag{7.30}$$

where

$$\begin{aligned} K_j(X_n, \theta) &= [\cos(u_j X_n) - \text{Re}\phi_{\theta}(u_j)] \frac{\partial \text{Re}\phi_{\theta}(u_j)}{\partial \theta} \\ &\quad + [\sin(u_j X_n) - \text{Im}\phi_{\theta}(u_j)] \frac{\partial \text{Im}\phi_{\theta}(u_j)}{\partial \theta} \end{aligned} \tag{7.31}$$

which is a vector of bounded functions of X_n with bounded first derivatives, hence, asymptotic normality of $\frac{\partial Q_N(\theta)}{\partial \theta}$ will follow by combining Lemma 2 and Proposition 1 if we show that

$$\lim_{N \rightarrow \infty} \text{E} \left[\sum_{n=1}^N \sum_{n'=1}^N K(X_n, \theta) K(X_{n'}, \theta)' \right] \tag{7.32}$$

is $O(N)$ if $X_n, n = 1, \dots, N$ is a Sup OU process, while it is $O(N^{2H}L(N))$ if $X_n, n = 1, \dots, N$ is a LRD Sup OU process. Here $K(X_n, \theta) = \sum_{j=1}^q K_j(X_n, \theta)$. In fact, asymptotic Normality of S_N , implied by Proposition 1, assures that the cumulants of S_N of order higher than 2 vanish asymptotically and these, by Lemma 2, dominate those of $\frac{\partial Q_N(\theta)}{\partial \theta}$. The result follows if the normalizing terms of S_N keep in focus the variance of the estimators $Q_N(\theta)$.

In order to evaluate (7.32), note that, from (7.31), apart from constant terms, we need to evaluate terms like $\text{Cov}(\cos(uX_n), \cos(vX_{n'}))$, $\text{Cov}(\cos(uX_n), \sin(vX_{n'}))$, $\text{Cov}(\sin(uX_n), \sin(vX_{n'}))$.

The above covariances may be conveniently tackled by considering well known trigonometric identities. By defining $\psi_{n,n'}(u, v) = \text{E}(e^{iuX_n + ivX_{n'}})$ we may estimate

$$\begin{aligned} & \text{Cov}(\cos(uX_n), \cos(vX_{n'})) = \\ & \frac{1}{4} [\psi_{n,n'}(u, v) + \psi_{n,n'}(u, -v) + \psi_{n,n'}(-u, v) + \psi_{n,n'}(-u, -v)] \\ & \quad - \frac{1}{4} [\phi(u) + \phi(-u)][\phi(v) + \phi(-v)] \end{aligned} \tag{7.33}$$

The above expression may be evaluated with the help of Lemma 2. We distinguish the two cases of X_n being Sup OU or LRD Sup OU.

If $X_n, n = 1, \dots, N$ is an LRD Sup OU process, from (7.18), for $n > n'$, the above expression, if we set $\exp\{uL_1[|n - n'|]|n - n'|^{-2(1-H)}\} = e_{n,n'}(u)$, after some algebra can be evaluated as

$$\begin{aligned} & \frac{1}{4} [(\phi(u)\phi(v) + \phi(u)\phi(-v))(e_{n,n'}(u) - 1) + (\phi(-u)\phi(v) + \\ & \quad + \phi(-u)\phi(-v))(e_{n,n'}(-u) - 1)] \end{aligned} \tag{7.34}$$

Next note that $e_{n,n'}(u) - 1 \sim |n - n'|^{-2(1-H)}$ as $|n - n'| \rightarrow \infty$. It is clear that all other terms appearing in (7.32) can be treated analogously and are of the same order, hence, using standard calculations we have that the variance of $\frac{\partial Q_N(\theta)}{\partial \theta}$ is $O(N^{2H}L(N))$ when $X_n, n = 1, \dots, N$ is a LRD Sup OU process.

If $X_n, n = 1, \dots, N$ is a Sup OU process, we may resort to expression (7.25) with a finite M instead of ∞ in the sum. The convergence to 0 in such a case is much quicker, assuring that the variance of $\frac{\partial Q_N(\theta)}{\partial \theta}$ is $O(N)$ when $X_n, n = 1, \dots, N$ is a LRD Sup OU process.

The second step of the proof is to show that the matrix $\frac{\partial^2 Q_N(\theta)}{\partial \theta \partial \theta'}$ converges to a fixed quantity for θ in a neighborhood of θ_0 . For $|\theta_0 - \hat{\theta}_N| \leq |\theta_0 - \hat{\theta}_N|$, we have

$$\begin{aligned} \frac{\partial^2 Q_N(\hat{\theta}_N)}{\partial \theta \partial \theta'} &= -2 \sum_{j=1}^q \frac{\partial \phi_{\hat{\theta}}(u_j)}{\partial \theta} \frac{\partial \phi_{\hat{\theta}}(u_j)}{\partial \theta'} + \\ &+ 2 \sum_{j=1}^q [\operatorname{Re} \phi_N(u_j) - \operatorname{Re} \phi_{\hat{\theta}}(u_j)] \frac{\partial^2 \operatorname{Re} \phi_{\hat{\theta}}(u_j)}{\partial \theta \partial \theta'} + \\ &+ 2 \sum_{j=1}^q [\operatorname{Im} \phi_N(u_j) - \operatorname{Im} \phi_{\hat{\theta}}(u_j)] \frac{\partial^2 \operatorname{Im} \phi_{\hat{\theta}}(u_j)}{\partial \theta \partial \theta'} \end{aligned} \quad (7.35)$$

Again, by assumptions of Theorem 1 and standard arguments we have that the r.h.s. of the above expression converges to $-2B(\theta_0)$ in probability. \square

References

- [1] Anh, V. V., Leonenko, N. N., Shieh, N. R., Taufer, E. (2010). Simulation of multifractal products of Ornstein-Uhlenbeck type processes. *Nonlinearity*, 23(4), 823.
- [2] Baran, S., Pap, G., van Zuijlen, M. C. A. (2003). Estimation of the mean of stationary and nonstationary Ornstein-Uhlenbeck processes and sheets. *Comput. Math. Appl.* 45, 563-579.
- [3] Barndorff-Nielsen, O.E. (1997). Normal inverse Gaussian distributions and stochastic volatility modelling. *Scand. J. Statist.* 24, no. 1, 1-13.
- [4] Barndorff-Nielsen, O.E. (1998). Processes of Normal Inverse Gaussian type. *Finance and Stochastics* 2, 41-68.
- [5] Barndorff-Nielsen, O. E. (2001). Superposition of Ornstein-Uhlenbeck type processes. *Theory Probab. Appl.* 45, no. 2, 175-194. Translated from *Teor. Veroyatnost. i Primenen.* 45 (2000), no. 2, 289-311.
- [6] Barndorff-Nielsen, O.E., Jensen, J.L., Sørensen, M. (1998). Some stationary processes in discrete and continuous time. *Adv. Appl. Prob.* 30 (4), 989-1007.
- [7] Barndorff-Nielsen, O.E., Shephard, N. (2001). Non-Gaussian Ornstein-Uhlenbeck-based models and some of their uses in financial economics. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 63, no. 2, 167-241.
- [8] Cox, D.R., (1991). Long-range dependence, nonlinearity and time irreversibility. *J. Time Ser. Anal.* 12, no. 4, 329-335.
- [9] De Bona, G. (2005). *Long memory e mercati azionari: uno studio empirico del MIB30*. Laurea thesis, University of Trento.
- [10] Ding, Z., Granger, C.W.J. (1996). Modeling volatility persistence of speculative returns: a new approach. *J. Econometrics* 73, no. 1, 185-215.
- [11] Dufresne, D. (1997). Algebraic properties of Beta and Gamma distributions, and applications. *Adv. Appl. Math.* 20, 285-299.
- [12] Feuerverger, A. and McDunnough, P. (1981). On some fourier methods for inference. *J. Amer. Stat. Assoc.* 76, 379-387.
- [13] Florens-Landais, D., Pham, H. (1999). Large deviations in estimation of an Ornstein-Uhlenbeck model. *J. Appl. Probab.* 36, 60-77.
- [14] Gloter, A. (2001). Parameter estimation for a discrete sampling of an integrated Ornstein-Uhlenbeck process. *Statistics* 35, 225-243.
- [15] Granger, C.W.J. (1980). Long memory relationships and the aggregation of dynamic models. *J. Econometrics* 14, no. 2, 227-238.
- [16] Grigelionis, B. (1999). Processes of Meixner type. *Liet. Mat. Rink.* 39, no. 1, 40-51; translation in *Lithuanian Math. J.* 39 (1999), no. 1, 33-41.
- [17] Knight, J.L. and Satchell, S.E. (1997). The cumulant generating function estimation method. *Econometric Theory* 13, 170-184.
- [18] Knight, J.L. and Yu, J. (2002). Empirical characteristic function in time series estimation. *Econometric Theory* 18, 691-721.
- [19] Kotz, S., Kozubowski, T. J., Podgórski, K. (2001). *The Laplace Distributions and Generalizations*. Birkhäuser, Boston.
- [20] Leonenko, N., Petherick, S., Taufer, E. (2013). Multifractal models via products of geometric OU-processes: Review and applications. *Physica A: Statistical Mechanics and its Applications*, 392(1), 7-16.
- [21] Leonenko, N.N., Taufer, E. (2005). Convergence of integrated superpositions of Ornstein-Uhlenbeck processes to fractional Brownian motion. *Stochastics* 77 (6), 477-499.
- [22] Leonenko, N., Taufer, E. (2013). Disaggregation of spatial autoregressive processes. *Spatial Statistics*, 3, 1-20.
- [23] Madan, D.B. and Seneta, E. (1987). Simulation of estimates using the empirical characteristic function *International Statistical Review* 55, 153-161.
- [24] Madan, D.B., Seneta, E. (1990). The VG model for share market returns. *J. Bus.* 63, 511-524.
- [25] Pap, G., van Zuijlen, M. C. A. (1996). Parameter estimation with exact distribution for multidimensional Ornstein-Uhlenbeck processes. *J. Multiv. Anal.* 59, 153-165.
- [26] Schoutens, W., Teugels, J.L. (1998). Lévy processes, polynomials and martingales. Special issue in honor of Marcel F. Neuts. *Comm. Statist. Stochastic Models* 14, no. 1-2, 335-349.
- [27] Schoutens, W. (2003). *Lévy Processes in Finance*. Wiley, Chichester.
- [28] Seneta, E. (2004) Fitting of Variance-Gamma model to financial data. *J. Appl. Probab.* 41 A, 177-187.
- [29] Steutel, F. W., van Harn, K. (2004). *Infinite Divisibility of Probability Distributions on the Real Line*. Marcel Dekker, New York.
- [30] Taufer, E., Leonenko, N., Bee, M. (2011). Characteristic function estimation of Stochastic volatility models. *Computational Statistics and Data Analysis* 55, 2525-2539.
- [31] Taufer, E., Leonenko, N. (2009). Characteristic function estimation of non-Gaussian Ornstein-Uhlenbeck processes. *Journal of Statistical Planning and Inference* 139, 3050-3063.
- [32] Taufer, E., Leonenko, N. (2009b). Simulation of Lévy-driven Ornstein-Uhlenbeck processes with given marginal distribution. *Computational Statistics and Data Analysis*, 53, 2427-2437.
- [33] Woerner, J.H.C. (2004) Estimating the skewness in discretely observed Lévy processes. *Econometric Theory* 20, 927-942.
- [34] Wolfe, J. (1982). On a continuous analogue of the stochastic difference equation $X_n = \rho X_{n-1} + B_n$. *Stochastic Process. Appl.* 12, no. 3, 301-312.