Poisson Driven Stationary Markov Models

Michelle Anzarut[1], Ramsés H. Mena[1], Consuelo R. Nava[2] and Igor Prünster[3]

[1] IIMAS, Universidad Nacional Autónoma de México, Mexico.
E-mail: michelle@sigma.iimas.unam.mx; ramses@sigma.iimas.unam.mx

[2] Department of Economics and Statistics and Collegio Carlo Alberto, Università degli Studi di Torino, Italy.
E-mail: consuelo.nava@carloalberto.org

[3] Department of Decision Sciences, BIDSA and IGIER, Bocconi University, Italy.
E-mail: igor@unibocconi.it

Abstract

We propose a simple yet powerful method to construct strictly stationary Markovian models with given but arbitrary invariant distributions. The idea is based on a Poisson-type transform modulating the dependence structure in the model. An appealing feature of our approach is the possibility to control the underlying transition probabilities and, therefore, incorporate them within standard estimation methods. Given the resulting representation of the transition density, a Gibbs sampler algorithm based on the slice method is proposed and implemented. In the discrete-time case, special attention is placed to the class of generalized inverse Gaussian distributions. In the continuous case, we first provide a brief treatment of the class of gamma distributions, and then extend it to cover other invariant distributions, such as the generalized extreme value class. The proposed approach and estimation algorithm are illustrated with real financial datasets.

Keywords: Bayesian inference; Generalized Extreme Value distribution; Generalized Inverse Gaussian distribution; Gibbs sampler; Markov process; Slice method; Stationary model.

1 Introduction

Stationarity and other stability properties represent a crucial component in the theory and application of stochastic processes. Indeed, in several modeling contexts, the assumption that some distributional features remain invariant over time is often needed to implement estimation and prediction procedures, or simply to be able to analytically determine quantities of interest. For example, within the classical theory of time series (Brockwell and

Davis; 1987), one looks for causality and invertibility conditions required by some estimation techniques; in mathematical finance many volatility and short–term interest rate models, leading to closed–form expressions of derivative pricing formulas are stationary (Linetsky; 2007); several of the tractable stochastic models in theoretical population genetics are stationary (Ewens; 2004) or have certain stability features; etc. If at the outset one starts with a model that allows for non-stationarity, often data transformations or parametric restrictions are needed in order to attain stability properties or even stationarity itself. However, these operations or restrictions are not always easy to attain or to handle in practice.

An alternative approach, undertaken in this paper, is to consider directly stationary models that accommodate empirical observations or a given phenomenon under study. With this premise, a natural starting point is to focus on Markov processes, and to look for transition mechanisms that retain a particular distribution of interest invariant over time. In fact, some of the drawbacks typically accounting for the use of non-stationary models can be outmatched by increasing the flexibility of the invariant distribution.

Defining Markov models with prescribed invariant distributions poses a tradeoff between marginal and conditional properties, as one can have several models with different dependence structure while retaining the same stationary distribution. This issue can be handled, to some extent, with a particular context in mind, e.g. fulfilling certain continuity or dependency requirements. Indeed, this is the approach followed by some of the constructions available in the literature. Most of these constructions rely on a thinning argument, i.e. decomposing the random variable at time t as a thinned version of the immediate past plus an innovation term. For instance, in Barndorff-Nielsen and Shephard (2001), the property of self-decomposability is used to attain such thinning and to characterize a class of continuous-time stationary models termed Ornstein-Uhlenbeck type processes. A vast literature on models with discrete invariant distributions is reviewed in McKenzie (2003), where binomial and other kind of thinning operators are employed to define discrete-valued time series models with geometric, binomial, negative binomial and Poisson marginals. An approach which enriches such thinning operators is due to Joe (1996) and Jørgensen and Song (1996), where discrete-time Markov processes with invariant distributions in the convolution-closed infinitely divisible class are presented. One-dimensional stationary diffusion processes of the mean reverting type, with prescribed marginal distributions, are explored in Bibby et al. (2005) by specifying a particular form of the diffusion coefficient modulating the process.

The above constructions are very appealing in statistical applications but, in most cases, no general expression or representation for the transition probabilities is available, particularly in the continuous-time setting. In fact, often one starts from a stochastic equation describing the dynamics of a phenomena in time and, thus, look for analytic expressions for the corresponding transition probabilities, which are not always immediate. However, full control of the transition probabilities driving a Markov process is always a desirable feature as estimation, simulation and prediction procedures become accessible.

Within a similar setup Pitt et al. (2002) exploit the reversibility property characterizing Gibbs sampler Markov chains to build strictly stationary AR(1)-type models with any choice of marginal distribution. Their approach is very general and requires to make choices of dependence to accommodate the specific modeling needs. This not always results in manageable expressions for the underlying transition probability. Some exceptions, meeting particular dependency or distributional requirements, can be found in Mena (2001), Mena and Walker (2005, 2007a,b, 2009), and Contreras-Cristán et al. (2009).

Here we aim at constructing stationary Markov models with tractable transition probabilities and prescribed arbitrary invariant distributions supported on $\mathbb{R}_+$. This is achieved by exploiting a symmetry induced by a Poisson-type transform, which allows to attain the desired invariant distribution and is coupled with a high degree of analytical tractability. Moreover, the proposed models can be extended, by means of simple transformations, to processes with invariant distributions supported on $\mathbb{R}$ or other state-spaces, while preserving the appealing transition probability tractability.

In the discrete-time case, attention is focused on the class of Generalized Inverse Gaussian (GIG) distributions. The GIG class is very flexible, and allows to obtain various explicit results. In the continuous-time setup, we use the gamma distribution as basic building block and obtain, via suitable transformations, a richer class of diffusion processes with known transition density. This includes, for instance, diffusions with generalized extreme value (GEV) invariant distributions which, to the best of our knowledge, have not been derived before. In order to perform Bayesian estimation for such processes, we derive a Gibbs sampling algorithm, based on some slice sampler techniques. The algorithm is implemented in a simulation study, and in the analysis of three financial datasets.

In the next sections we proceed as follows. We devote Section 2 to the discrete–time case, we describe the general construction based on the Poisson style transform, some of its general properties, and study the GIG case in detail. In Section 3, we extend the model to the continuous-time setup by introducing a time-homogeneous dependence structure. With a suitable transformation of the gamma invariant distribution case, we obtain a wider class of diffusion models with known transition probability. We derive a Gibbs sampling algorithm for the estimation over Section 4, and illustrate it in synthetic and real data in Section 5. Finally, Section 6 contains some concluding remarks and future research directions.

2 Poisson-driven Markov process

2.1 The construction in discrete time

We first provide a definition of the transformed density we will be using to drive the dependence in our construction.

Definition 1. Let f be an absolutely continuous probability density function supported on $\mathbb{R}_+$. For any $\phi > 0$ and $y \in \mathbb{N} \cup \{0\}$ we define the *Poisson weighted density* as

$$\hat{f}(x; y, \phi) := \frac{x^y e^{-x\phi} f(x)}{\xi(y, \phi)}, \tag{1}$$

where

$$\xi(y, \phi) := \int_{\mathbb{R}_+} z^y e^{-z\phi} f(z) \mathrm{d}z.$$

Note that (1) is well defined as $\xi(y, \phi)$ can be seen as a moment of an exponentially tilted positive random variable, which always exists for $\phi > 0$. Moreover, when $\phi \downarrow 0$, the Poisson weighted density reduces to the size-biased density of f and, when $y = 0$, it reduces to the Esscher transform of f. The density (1) can also be seen as the posterior density of a Poisson distribution with parameter ϕx, denoted as $\text{Po}(\phi x)$, with prior f on x. This latter aspect, combined with the general idea of weighted distributions introduced by Rao (1965), explains the name we have attributed to (1).

Moments and the Laplace transform of random variables with distribution (1) are readily available. Indeed, if X has density (1), then

$$\mathbb{E}[X^r] = \frac{\xi(y + r, \phi)}{\xi(y, \phi)} \qquad \text{and} \qquad \mathcal{L}_X(\lambda) = \frac{\xi(y, \phi + \lambda)}{\xi(y, \phi)}, \tag{2}$$

where $\mathcal{L}_X(\lambda) := \mathbb{E}[e^{-\lambda X}]$.

Based on the above Poisson weighted density, we can construct a stationary Markov process $(X_n)_{n\in\mathbb{Z}_+}$ with invariant distribution having density f. To this end, define the following time–homogeneous one–step ahead Markovian density

$$\begin{aligned} p(x_{n-1}, x_n) &:= \sum_{y=0}^{\infty} \hat{f}(x_n; y, \phi) \, \mathsf{Po}(y; x_{n-1}\phi) \\ &= \exp\{-\phi(x_n + x_{n-1})\} f(x_n) \sum_{y=0}^{\infty} \frac{(x_n x_{n-1}\phi)^y}{y! \xi(y, \phi)}, \end{aligned} \tag{3}$$

which clearly satisfies the detailed balance condition

$$p(x_{n-1}, x_n) f(x_{n-1}) = p(x_n, x_{n-1}) f(x_n)$$

for all $x_{n-1}, x_n \in \mathbb{R}_+$, leading to a time-reversible Markov process. It is immediate to verify that f is invariant under transition (3), meaning

$$\int_{\mathbb{R}_+} p(x_{n-1}, x_n) f(x_{n-1}) \mathrm{d}x_{n-1} = f(x_n).$$

Therefore, the process driven by (3) is strictly stationary.

Definition 2. The stationary Markov process, driven by transition density (3) and with stationary density f, is termed *f-stationary Poisson-driven Markov process.*

As far as conditional moments $\mathbb{E}_{x_{n-1}}[X_n^r] := \mathbb{E}[X_n^r \mid X_{n-1} = x_{n-1}]$ of an f-stationary Poisson-driven Markov process are concerned, the combination of (2) and (3) leads to

$$\mathbb{E}_{x_{n-1}}[X_n^r] = \sum_{y=0}^{\infty} \left[\frac{\xi(y+r,\phi)}{\xi(y,\phi)}\right] \mathsf{Po}(y; x_{n-1}\phi).$$

Consequently, provided that f admits second moment, the autocorrelation can be expressed as

$$\mathsf{Corr}(X_n, X_{n-1}) = \frac{1}{\sigma_f^2}\left[\sum_{y=0}^{\infty}\left(\frac{\xi(y+1,\phi)^2}{\xi(y,\phi)}\right)\frac{\phi^y}{y!} - \mu_f^2\right],$$

where μ_f and σ_f^2 denote the mean and variance of the stationary density. For example if f is chosen to be the density of a $\mathsf{Ga}(a,b)$ distribution, then the above correlation reduces to

$$\mathsf{Corr}(X_n, X_{n-1}) = \frac{\phi}{b+\phi}.$$

Once the form of f is chosen, the dependence in the model is driven by the parameter ϕ. See, for example, Figure 1. In particular, when ϕ goes to infinity the correlation tends to one. In the following section we focus on a general and flexible class of densities on $\mathbb{R}_+$.

2.2 GIG-stationary Poisson-driven Markov process

Part of the dependence in the model is induced by the choice of marginal density, f, which in turn can be selected by the nature of the phenomenon or data under study. The other part is due to the dependence parameter ϕ. In this section we focus on densities f belonging to the family of generalized inverse Gaussian (GIG) distributions, with density

$$\mathsf{GIG}(x;\alpha,\delta,\gamma) = \frac{1}{A(\alpha,\delta,\gamma)} x^{\alpha-1} \exp\left\{-\frac{1}{2}(\delta^2 x^{-1} + \gamma^2 x)\right\} \mathbb{I}_{\{x>0\}},$$

where $A(\alpha,\delta,\gamma) := (\delta/\gamma)^{\alpha}\, 2K_{\alpha}(\delta\gamma)$ and K_ν denotes the modified Bessel function of the third type with index ν. The parameter domain is $\alpha \in \mathbb{R}$, $(\delta,\gamma) \in \Theta_\alpha$ with

$$\Theta_\alpha = \begin{cases} \delta \geq 0, \gamma > 0 & \text{if } \alpha > 0, \\ \delta > 0, \gamma > 0, & \text{if } \alpha = 0, \\ \delta > 0, \gamma \geq 0, & \text{if } \alpha < 0. \end{cases}$$

The values $\delta = 0$ and $\gamma = 0$ are interpreted as limiting cases. Some well known distributions are particular cases of the GIG family, e.g. gamma ($\alpha > 0, \delta = 0, \gamma > 0$), inverse gamma ($\alpha > 0, \delta > 0, \gamma = 0$), and inverse Gaussian ($\alpha = -\frac{1}{2}, \gamma > 0, \delta > 0$). Indeed, the class of GIG distributions appears frequently in various fields of applications. See Eberlein and Hammerstein (2004) for an account on GIG distributions. Recently they gained popularity also within Bayesian contexts as key ingredients to build some general distributions on the simplex. See, e.g., Favaro et al. (2011, 2012).

Developing further the construction of Section 2.1 for f being a GIG density, we obtain

$$\xi(y, \phi) = \frac{A(\alpha + y, \delta, \sqrt{\gamma^2 + 2\phi})}{A(\alpha, \delta, \gamma)}.$$

It is then straightforward to see that the Poisson weighted distribution generated by a GIG density f is also GIG, i.e.

$$\hat{f}(x; y, \phi) = \mathsf{GIG}(x; \alpha + y, \delta, \sqrt{\gamma^2 + 2\phi}),$$

which means that the GIG family is closed under Poisson weighted transformations, an appealing feature for simulation and estimation purposes. Given this, the corresponding transition density (3) is then of the form

$$\begin{aligned} p(x_{n-1}, x_n) &= \sum_{y=0}^{\infty} \mathsf{GIG}(x_n; \alpha + y, \delta, \sqrt{\gamma^2 + 2\phi}) \, \mathsf{Po}(y; x_{n-1}\phi) \\ &= x_n^{\alpha-1} \exp\left\{-\phi(x_n + x_{n-1}) - \frac{1}{2}\left[\frac{\delta^2}{x_n} + \gamma^2 x_n\right]\right\} \sum_{y=0}^{\infty} \frac{(x_{n-1}x_n\, \phi)^y}{y!\; A(\alpha + y, \delta, \sqrt{\gamma^2 + 2\phi})}. \end{aligned}$$

Some particular cases offering further simplifications are at hand:

• IG–*stationary Poisson-driven Markov process.* For $\alpha = -\frac{1}{2}$, we obtain the inverse Gaussian distribution, i.e. $\mathsf{IG}(\delta, \gamma) = \mathsf{GIG}(-1/2, \delta, \gamma)$, leading to

$$\xi(y, \phi) = \frac{A(y - \frac{1}{2}, \delta, \sqrt{\gamma^2 + 2\phi})}{A(-\frac{1}{2}, \delta, \gamma)} = \sqrt{\frac{2}{\pi}}\, e^{\delta\gamma}\, \delta^{y+\frac{1}{2}} \left(\sqrt{\gamma^2 + 2\phi}\right)^{\frac{1}{2}-y} K_{y-\frac{1}{2}}\left(\delta\sqrt{2\phi + \gamma^2}\right),$$

and transition density

$$\begin{aligned} p(x_{n-1}, x_n) = \exp&\left\{-\phi(x_{n-1} + x_n) - \frac{1}{2}\left[\delta^2 x_n^{-1} + \gamma^2 x_n\right]\right\} x_n^{-\frac{3}{2}} \left(\frac{\delta}{\sqrt{\gamma^2 + 2\phi}}\right)^{\frac{1}{2}} \\ &\sum_{y=0}^{\infty} \frac{\left(\phi\; x_{n-1}\; x_n \sqrt{\gamma^2 + 2\phi}\; \delta^{-1}\right)^y}{y!\; 2K_{y-\frac{1}{2}}(\delta\sqrt{\gamma^2 + 2\phi})}. \end{aligned}$$

• Ga-*stationary Poisson-driven Markov process.* For $\delta = 0$, the gamma distribution is recovered, i.e. $\mathsf{Ga}(\alpha, \beta) = \mathsf{GIG}(\alpha, 0, \gamma)$ where $\beta = \gamma^2/2$ resulting in

$$\xi(y, \phi) = \frac{\beta^\alpha}{\Gamma(\alpha)} \frac{\Gamma(\alpha + y)}{(\beta + \phi)^{y+\alpha}},$$

with corresponding transition density

$$p(x_{n-1},x_n) = \frac{\exp\{-[\phi(x_n+x_{n-1})+\beta x_n]\}}{(\phi+\beta)^{-(\alpha+1)/2}\,\phi^{(\alpha-1)/2}} \left(\sqrt{\frac{x_n}{x_{n-1}}}\right)^{\alpha-1} \mathrm{I}_{\alpha-1}\left(2\sqrt{x_{n-1}x_n\phi(\phi+\beta)}\right),$$

where $\mathrm{I}_\nu(x)$ denotes the modified Bessel function of the first kind.

• IGa-*stationary Poisson-driven Markov process.* For $\gamma=0$, a GIG random variable reduces to the inverse gamma distribution, i.e. $\mathsf{GIG}(-\alpha,\delta,0)=\mathsf{IGa}(a,b)$ with $a=-\alpha, b=\frac{\delta^2}{2}$. This leads to

$$\xi(y,\phi) = 2\,\frac{b^{\frac{a+y}{2}}\,\phi^{\frac{a-y}{2}}}{\Gamma(a)}\,K_{y-a}(2\sqrt{b\,\phi}),$$

and

$$p(x_{n-1},x_n) = x_n^{-a-1}\exp\left\{-\phi(x_n+x_{n-1})-\frac{b}{x_n}\right\}\left(\frac{b}{\phi}\right)^{\frac{a}{2}}\sum_{y=0}^{\infty}\frac{\left(\phi\,x_{n-1}\,x_n\,\sqrt{\frac{\phi}{b}}\right)^y}{y!2\,K_{y-a}(2\sqrt{b\,\phi})}.$$

The particular case of $a=1/2$ corresponds to the positive $\frac{1}{2}$–stable distribution.

3 Extension to continuous time

3.1 The general case

A natural question that arises from the definition of f-stationary Poisson-driven Markov process is whether it can be extended to a continuous-time processes. The crucial point for deriving such an extension is to verify the resulting transition density (3) also satisfies the Chapman-Kolmogorov equation (see Karlin and Taylor; 1981), assuring the Markov property remains valid. This, clearly requires assuming additional conditions, e.g. to specify the behavior at any infinitesimal time. A typical way to achieve this is to consider the discrete-time process as embedded into a continuous-time process, and then perform a suitable time transformation to obtain a representation of the continuous process in the limit. Alternatively, one could let the underlying time-homogeneous effect enters through one of the dependency parameters in (3), in a way that the Chapman-Kolmogorov equation still remains valid. Here we follow the latter approach and consider time homogeneous transition densities

$$p_t(x_0,x_t) = \exp\{-\phi_t(x_0+x_t)\}\;f(x_t)\sum_{y=0}^{\infty}\frac{(x_t x_0\phi_t)^y}{y!\xi(y,\phi_t)}, \tag{4}$$

with $t\mapsto\phi_t$ a continuous function to be chosen such that the Chapman-Kolmogorov equation is satisfied. In terms of Laplace transforms, we need to find the functional form of ϕ_t, such that

$$\mathcal{L}_{X_{t+s}|X_0}(\lambda) = \mathbb{E}_{X_0}[\mathcal{L}_{X_{t+s}|X_s}(\lambda)], \tag{5}$$

where the Laplace transform of the corresponding transition distribution is given by

$$\mathcal{L}_{X_t|X_0}(\lambda) = \sum_{y=0}^{\infty} \mathsf{Po}(y; x_0\phi_t) \frac{\xi(y, \phi_t + \lambda)}{\xi(y, \phi_t)}.$$

Therefore, provided ϕ_t satisfies (5), one can use (4) to define a continuous-time Markov process.

3.2 Continuous Ga-stationary Poisson-driven Markov process

Let us focus on a particular case of the GIG-stationary Poisson-driven Markov process, namely the case of the gamma invariant distribution. The appealing feature of this case is the availability of the explicit functional form for ϕ_t that assures the Chapman-Kolmogorov equation is satisfied. Moreover, as it will be shown in Section 3.3, it represents a gateway for the construction of other continuous models.

Indeed, as seen in Section 2.2, for the specific case of the Ga-stationary Markov process we have

$$\xi^{\mathsf{Ga}}(y, \phi) = \frac{b^a}{\Gamma(a)} \frac{\Gamma(a+y)}{(b+\phi)^{y+a}} \qquad \text{and} \qquad \hat{f}^{\mathsf{Ga}}(x; y, \phi) = \mathsf{Ga}(x; y+a, b+\phi)$$

which, within a continuous-time framework, leads to a transition density of the form

$$p_t^{\mathsf{Ga}}(x_0, x_t) = \frac{\exp\{-[\phi_t(x_t + x_0) + bx_t]\}}{(\phi_t + b)^{-(a+1)/2}\, \phi_t^{(a-1)/2}} \sqrt{\frac{x_t}{x_0}}^{\,a-1} \mathrm{I}_{a-1}\left(2\sqrt{x_t x_0 \phi_t(\phi_t + b)}\right). \tag{6}$$

It can be shown (Mena and Walker; 2009) that for such a transition density the Chapman-Kolmogorov equation is satisfied if and only if

$$\phi_t = \frac{b}{e^{ct} - 1}, \qquad c > 0. \tag{7}$$

Given the continuous time and space nature of this process, there are two options for this Markov process $(X_t)_{t\geq 0}$, either it corresponds to the law of a diffusion process, or it corresponds to the law of a continuous-time jump process. This would give us a full stochastic characterization of the model and its path continuity properties. In particular, it is well-known (Karlin and Taylor; 1981) that if

$$\lim_{h\downarrow 0} \frac{1}{h} \mathbb{E}[|\Delta_h X_t|^p \mid X_t = x_t] = 0 \quad \text{for} \quad p > 2, \quad \text{and} \quad \Delta_h X_t := X_{t+h} - X_t, \tag{8}$$

then $(X_t)_{t\geq 0}$ cannot have jump discontinuities and, thus, results in a diffusion process characterized by its infinitesimal conditional mean and variance coefficients

$$\mu(x,t) = \lim_{h\downarrow 0} \frac{1}{h} \mathbb{E}_x[\Delta_h X_t] \qquad \text{and} \qquad \sigma^2(x,t) = \lim_{h\downarrow 0} \frac{1}{h} \mathbb{E}_x[(\Delta_h X_t)^2]. \tag{9}$$

Hence, by verifying condition (8), it follows that $(X_t)_{t\geq 0}$ is indeed a diffusion process and, by computing the limits (9), $(X_t)_{t\geq 0}$ can be seen as the solution of

$$\mathrm{d}X_t = c\left(\frac{a}{b} - X_t\right)\,\mathrm{d}t + \sqrt{\frac{2\,c}{b}\,X_t}\,\mathrm{d}W_t, \tag{10}$$

where $(W_t)_{t\geq 0}$ denotes a standard Brownian motion. Note that (10), in turn, represents a reparametrization of the mean reverting Cox-Ingersoll-Ross (CIR) model (Cox et al.; 1985), commonly used to model nominal interest rates.

Although, expression (6) for the transition density is well known in the literature, representation (3) turns out to be an attractive alternative. This latter expression is particularly useful from the perspective of simulation and estimation of a CIR model, since it only involves gamma and Poisson distributions which are straightforward to handle computationally. Moreover, by using this model we are able to build up other models by a simple transformation as will be shown in the next section.

3.3 Models derived from a Ga-Stationary Poisson driven Markov process

Given the previous construction, with a simple transformation we can obtain other classes of continuous-time Markov processes. In particular, let us assume that we want to construct an f-stationary Poisson-driven Markov process, where f is the density corresponding to a random variable $X = h(Z)$ with $Z \sim \mathsf{Ga}(a, b)$, and h is a $\mathbb{R}$-valued function with known and differentiable inverse. Notice that this implies that the model can be extended to cover marginals f with support $\mathbb{R}$.

Letting $g(x) := h^{-1}(x)$ and $\mathcal{J}(x) := |g'(x)|$, we have that $f(x) = \mathsf{Ga}\,(g(x); a, b)\,\mathcal{J}(x)$, and the corresponding transition probability is given by

$$p_t(x_0, x_t) \quad = \quad \sum_{y=0}^{\infty} \hat{f}(x_t; y, \phi_t)\,\mathsf{Po}(y; g(x_0)\phi_t). \tag{11}$$

Also, it is easily verified that

$$\xi(y, \phi, g) = \xi^{\mathsf{Ga}}(y, \phi) \quad \text{ and } \quad \hat{f}(x; y, \phi) = \hat{f}^{\mathsf{Ga}}(g(x); y, \phi)\mathcal{J}(x), \tag{12}$$

and the transition (11) can then be simplified as

$$\begin{aligned} p_t(x_0, x_t) &= e^{-\phi_t(g(x_0)+g(x_t))} f(x_t) \sum_{y=0}^{\infty} \frac{[g(x_t)\,g(x_0)\phi_t]^y}{y!\xi(y, \phi_t, g)} \\ &= p_t^{\mathsf{Ga}}(g(x_0), g(x_t))\mathcal{J}(x_t), \end{aligned}$$

leading to a large class of tractable continuous f-stationary Poisson-driven Markov processes $(X_t)_{t\geq 0}$.

If the transformation h is also twice differentiable, then by applying Itô's lemma to (10), the associated transformed diffusion can be seen as solution to

$$\mathrm{d}X_t = c\left[h'\left(g(X_t)\right)\frac{a-b\,g(X_t)}{b} + h''\left(g(X_t)\right)\frac{g(X_t)}{b}\right]\mathrm{d}t + h'\left(g(X_t)\right)\sqrt{\frac{2\,c}{b}g(X_t)}\,\mathrm{d}W_t.$$

Here, the key aspect to remark, which represents a highly attractive feature in terms of practical implementation, is that we still have representation (11) for the transition density with the same ϕ_t function as in (7). Furthermore, this derivation allows us to avoid the challenging task of having to solve the Chapman-Kolmogorov equation directly for each case. For details on the above derivations see Appendix A.1 (cf. Supporting information).

To make things more concrete let us consider the case of generalized extreme value distributions (GEV). A GEV distribution is characterized by three parameters representing location $\mu \in \mathbb{R}$, scale $\sigma > 0$ and shape $\nu \in \mathbb{R}$. In symbols, $X \sim \mathsf{GEV}(\mu, \sigma, \nu)$. Its support depends on the value of ν, if $\nu = 0$ the support coincides with $\mathbb{R}$, whereas if $\nu > 0$ or $\nu < 0$ it coincides, respectively, with $(\mu - \sigma/\nu, \infty)$ and $(-\infty, \mu - \sigma/\nu)$. These three cases are often referred to as type I, II and III GEV distributions, a terminology which will also be adopted here. The cumulative distribution function associated to a GEV random variable is of the form

$$F(x;\mu,\sigma,\nu) = \exp\left\{-\left(1+\nu\frac{x-\mu}{\sigma}\right)^{-1/\nu}\right\},$$

where the case $\nu = 0$ is to be understood in the limiting sense. As far as its relation with other well–known distributions is concerned, a type I GEV coincides with a Gumbel distribution when $\nu = 0$, whereas from a type II GEV one immediately obtains a Fréchet distribution by setting $\nu = 1/\alpha > 0$, $\mu = 0$ and applying the change of variable $y = \sigma + x/\alpha$. From the type III GEV, one obtains a Weibull distribution by setting $\nu = -1/\alpha < 0$, $\mu = 0$ and applying the change of variable $y = \sigma - x/\alpha$.

Identifying a suitable choice of the function h, we can derive a stationary Poisson–driven Markov process displaying any of the above distributions (and others) as its stationary distribution. As an example, in Appendix A.2 (cf. Supporting information) we present two tables. The first one shows the transform function h, and the resulting conditional distributions for the inverse gamma, the Pareto, the GEV, the Gumbel, the Fréchet, and the Weibull distributions. The second table contains the corresponding stochastic differential equations (SDEs) for each case.

Remark 1. The parametric family of GEV distributions represents a major tool in modeling and measuring events with low probability of appearance. In particular, this family is widely used in Economics, Finance, Insurance and Risk Management, among other disciplines. See, for example, Klüppelberg (2002), Embrechts et al. (1997) and Coles (2001).

In many studies involving GEV distributions, the assumption of time-independent observations is made, which, however, is often not satisfied due to the persistence of extreme conditions over consecutive observations. An approach to overcome this issue, under stationarity assumption, is the one proposed by Coles (2001), changes across time in maxima are modeled via GEV distributions with parameters specified in terms of polynomials on time. In a similar direction, Huerta and Sansó (2007) propose, instead of polynomials, to use dynamic linear models. A related, but different, idea is the one proposed by Nakajima et al. (2012), where innovations of the underlying latent process follow a type I GEV distribution and induce an observation response again characterized by a type I GEV distribution. Although some of these approaches are similar in spirit to the one set forth here, it is important to remark that none of them undertakes the modeling from a continuous-time perspective. Indeed, to the best of our knowledge, the diffusion models derived from the GEV-Stationary Poisson driven process construction (Appendix A.2 in Supporting information) are unique in this sense. Namely, as diffusion models with GEV-invariant distributions and known transition density. Noteworthy examples, concerning series derived from the S&P 500 and the Tokyo Stock Exchange Price Index, are analyzed in Section 5.2 by resorting to such models.

4 Bayesian Estimation

The availability of a tractable expression for the transition density is highly desirable in the analysis and estimation of Markov processes. In this section we focus on estimation in the continuous-time case, since the discrete-time case can be easily recovered from it. First note that, if the choice of f leads to a manageable analytic expression in (3), the likelihood for a set of observations $\mathbf{x} = (x_1, \ldots, x_N)$ with $x_n := x_{t_n}$ is given by

$$\mathcal{L}_{\mathbf{x}}(\boldsymbol{\theta}) = f(x_1; \boldsymbol{\theta}) \prod_{n=2}^{N} p(x_{n-1}, x_n; \boldsymbol{\theta}),$$

where $\boldsymbol{\theta}$ denotes, generically, the set of parameters in the marginal f and the ones inherent to the dependency function ϕ_t. Alternatively, if the choice of f does not allow to perform the summation in (3) analytically, one could still make use of such a representation for the transition density and obtain an augmented version of the likelihood (Dempster et al.; 1997) via

$$\begin{aligned}\mathcal{L}^{aug}_{\mathbf{x},\mathbf{y}}(\boldsymbol{\theta}) =& f(x_1; \boldsymbol{\theta}) \prod_{n=2}^{N} \hat{f}(x_n; y_n, \phi_{\tau_n}, \boldsymbol{\theta}) \, \mathsf{Po}\left(y_n; \phi_{\tau_n} x_{n-1}, \boldsymbol{\theta}\right) \\ =& \exp\left\{-\sum_{n=2}^{N} \phi_{\tau_n}(x_n + x_{n-1})\right\} \left[\prod_{n=1}^{N} f(x_n; \boldsymbol{\theta})\right] \left[\prod_{n=2}^{N} \frac{(x_n x_{n-1} \phi_{\tau_n})^{y_n}}{y_n! \xi(y_n, \phi_{\tau_n}; \boldsymbol{\theta})}\right],\end{aligned}$$

where $\mathbf{y} = (y_2, \dots, y_N)$ is a vector of latent variables and ϕ_{τ_n} is the homogeneous time effect with $\tau_n = t_n - t_{n-1}$. The discrete-time case can be easily recovered when $t_n = n$ and $\tau_n = 1$ for every $n = 2, \dots, N$.

Given such a likelihood, the idea now consists in deriving an MCMC algorithm for Bayesian inference which incorporates it. In order to build a general estimation algorithm that does not rely on the availability of an analytical expression for (3), we propose a Gibbs sampler using some slice-technique ideas as those implemented in Kalli et al. (2011); Yau et al. (1965); Mena et al. (2011), Papaspiliopoulos and Roberts (2008). We can construct a Gibbs sampler algorithm based on an augmented representation of transition density (4), given by

$$p_t(x_0, x_t, u, y) = \frac{1}{\psi_y}\mathbb{I}(u < \psi_y)\exp\{-\phi_t(x_t + x_0)\}\, f(x_t; \boldsymbol{\theta})\, \frac{(x_t\, x_0\, \phi_t)^y}{y!\,\xi(y, \phi_t; \boldsymbol{\theta})},$$

where $y \mapsto \psi_y$ is a positive decreasing invertible function, termed truncation function, e.g. $\psi_y = e^{-\eta y}$, for $\eta > 0$. More specifically, with the latent variable u uniformly distributed given y, the augmented likelihood for a set of observations $\mathbf{x} = (x_1, \dots, x_N)$ at times $(t_1, \dots, t_N)$ can be simplified to

$$\begin{aligned}\mathcal{L}_{\mathbf{x},\mathbf{u},\mathbf{y}}(\boldsymbol{\theta}) = \exp\left\{-\sum_{n=2}^{N}\phi_{\tau_n}(x_n + x_{n-1})\right\}\left[\prod_{n=1}^{N} f(x_n; \boldsymbol{\theta})\right] \\ \times\left[\prod_{n=2}^{N}\frac{(x_n x_{n-1}\phi_{\tau_n})^{y_n}}{y_n!\,\xi(y_n, \phi_{\tau_n}; \boldsymbol{\theta})\psi_{y_n}}\,\mathbb{I}(u_n < \psi_{y_n})\right],\end{aligned}$$

leading to a log-likelihood

$$\begin{aligned} l_{\mathbf{x},\mathbf{u},\mathbf{y}}(\boldsymbol{\theta}) &= -\sum_{n=2}^{N}\phi_{\tau_n}(x_n + x_{n-1}) + \sum_{n=1}^{N}\log(f(x_n; \boldsymbol{\theta})) + \sum_{n=2}^{N}\left[\log(\mathbb{I}(u_n < \psi_{y_n})) - \log(\psi_{y_n})\right] \\ &+ \sum_{n=2}^{N}\left\{y_n \log(x_n x_{n-1}\phi_{\tau_n}) - \log(y_n!\,\xi(y_n, \phi_{\tau_n}; \boldsymbol{\theta}))\right\}, \end{aligned}$$

for $n = 2, \dots, N$, where $\tau_n := t_n - t_{n-1}$, $\mathbf{u} = (u_2, \dots, u_N)$ and $\mathbf{y} = (y_2, \dots, y_N)$.

Note that in the above expressions ϕ could also depend on $\boldsymbol{\theta}$. This expression is quite appealing in terms of the derivation of the full conditionals within the Gibbs sampler. In particular, if π denotes the prior distribution on $\boldsymbol{\theta}$, the corresponding full log-posterior distribution can be considerably simplified by separating the parameters in the stationary distribution, $\boldsymbol{\theta}^{(st)}$, and the parameters referring to the transition probability, but not to the stationary distribution, $\boldsymbol{\theta}^{(tr)}$. For example, in the case of the models in Section 3.3, derived from the $\mathsf{Ga}(a, b)$-Stationary Poisson driven processes, we have that $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(st)}, \boldsymbol{\theta}^{(tr)})$ with $\boldsymbol{\theta}^{(st)} = (a, b)$ and $\boldsymbol{\theta}^{(tr)} = c$. For the discrete-time case, we would have $\boldsymbol{\theta}^{(tr)} = \phi$.

The full log-posterior distribution for $\boldsymbol{\theta} = (\boldsymbol{\theta}^{(st)}, \boldsymbol{\theta}^{(tr)})$, under the assumption of independent prior distributions for each block of parameters, reduces to

$$\log \pi(\boldsymbol{\theta}^{(st)} \mid \cdots) \propto \log \pi(\boldsymbol{\theta}^{(st)}) + \sum_{n=1}^{N} \log(f(x_n; \boldsymbol{\theta}^{(st)})) - \sum_{n=2}^{N} \log(\xi(y_n, \phi_{\tau_n}; \boldsymbol{\theta})),$$

and

$$\log \pi(\boldsymbol{\theta}^{(tr)} \mid \cdots) \propto \log \pi(\boldsymbol{\theta}^{(tr)}) + \sum_{n=2}^{N} y_n \log(\phi_{\tau_n}) - \sum_{n=2}^{N} \phi_{\tau_n}(x_n + x_{n-1}) - \sum_{n=2}^{N} \log(\xi(y_n, \phi_{\tau_n}; \boldsymbol{\theta})).$$

Therefore, simulating from the full posteriors can be easily achieved, for instance, via the adaptive rejection Metropolis sampling (ARMS) algorithm. The full conditional distributions for the latent variables can be obtained componentwise via

$$\begin{aligned} \pi(u_n \mid \cdots) &= \mathsf{U}(u_n | 0, \psi(y_n)), \\ \pi(y_n \mid \cdots) &\propto \frac{[x_n\, x_{n-1}\, \phi_{\tau_n}]^{y_n}}{y_n!\, \xi(y_n, \phi_{\tau_n}; \boldsymbol{\theta})\, \psi_{y_n}} \mathbb{I}(u_n < \psi_{y_n}), \end{aligned} \tag{13}$$

for $n = 2, \ldots, N$. Here $\mathsf{U}(x|a, b)$ denotes a uniform distribution with parameters a and b. Note that the above distribution has support $y_n = 0, \ldots, \lfloor \psi^*(u_n) \rfloor$, where ψ^* denotes the inverse of ψ. This is precisely the advantage of using the slice mechanism, namely that we only need to sample from a finite support instead of a distribution supported on $\mathbb{N}$. Therefore, a relatively simple Gibbs sampling algorithm can be implemented with the above full conditional distributions. In Appendix B.1 (cf. Supporting information) details for their derivation are provided.

Remark 2. It is well-known that transition densities corresponding to stationary diffusion processes can be represented via spectral decompositions of the type

$$p_t(x_0, x_t) = f(x_t) \sum_{n=0}^{\infty} e^{-\lambda_n t} \varphi_n(x_0) \varphi_n(x_t),$$

with $\{\lambda_n\}_{n=0}^{\infty}$, $0 < \lambda_0 < \lambda_1 < \cdots$ eigenvalues and $\{\varphi_n(x)\}_{n=0}^{\infty}$ the corresponding eigenfunctions (Karlin and Taylor; 1981). However, the arguments in the above summation might be negative, thus making its efficient evaluation a challenging numerical problem, e.g. truncation of the summation becomes meaningless. In contrast, in our models the arguments in the summation of the corresponding transition density representation (4) are always positive. Therefore, deterministic or random truncations, such as the one implicitly performed when using the slice method in the proposed Gibbs sampler, are more reasonable and numerically efficient.

5 Illustrations

Over Section 5 we test the model and the estimation procedure. In Section 5.1, we do it by simulating Poisson-driven Markov processes with GIG and GEV stationary distributions in discrete and continuous time, respectively. In Section 5.2, we test the model and estimation over some financial datasets. First, we model the FTSE 100 equity index by means of a discrete-time GIG-stationary Poisson driven Markov process. Afterwards, we perform a continuous-time extreme value analysis by applying the model with GEV marginal density to the minimum daily stock returns of the S&P 500 and the Tokyo Stock Price Index. All the following results are based on 10000 iterations of the Gibbs sampler, with a burn–in of 1000 sweeps, and with a simulation every 50 kept.

5.1 Simulated Data

As mentioned above, in order to evaluate the performance of our approach, we test the model with two specific forms of stationary distributions, a GIG and a GEV. These choices are done given their relevance in areas such as Economics and Finance. See, e.g., Nakajima et al. (2012). The GIG case consists in two series of 1,500 simulated observations from a discrete-time model with $\mathsf{GIG}(\alpha, \delta, \gamma)$ invariant distribution, where $(\alpha, \delta, \gamma) = (1, 2, 3)$. In the first series $\phi = 1$, while in the second one $\phi = 80$. The two series paths appear in Figure 1. We notice that, as the dependence parameter ϕ gets larger, cluster structures appear in the data.

We perform the estimation method on the two data sets. In doing so, we assign independent unitary exponential priors to δ, γ, and ϕ, an independent standard normal prior for α, and then derive the full conditional distributions among with some other relevant quantities in Appendix B.2 (cf. Supporting information). Additionally, we set the truncation function ψ_y equal to $e^{-\eta y}$ with $\eta = 0.4$, this choice allows to have a known inverse ψ^* and, therefore, to immediately identify the support of the latent variable y.

To measure the accuracy of the stationary distribution approximation we will use the Kullback-Leibler divergence (KL-divergence) (see Kullback and Leibler; 1951). Such a divergence provides a global measure of the fit, avoiding a problem that often arises when inferring in the GIG family due to the fact that several parametrizations may lead to similar distributions. The KL-divergence between the real invariant distribution and the estimated one is 0.0011 for the first data series, and 0.0019 for the second one (see Figure 2). For the ϕ parameter, the posterior distribution mode is given by 1.299 in the first series, and 66.236 in the second one. Hence, there is clear evidence of the good performance of the proposed model.

An immediate procedure for obtaining a set of m trajectories in new times is available by, first, simulating m parameter values from the posterior distributions, and, second,

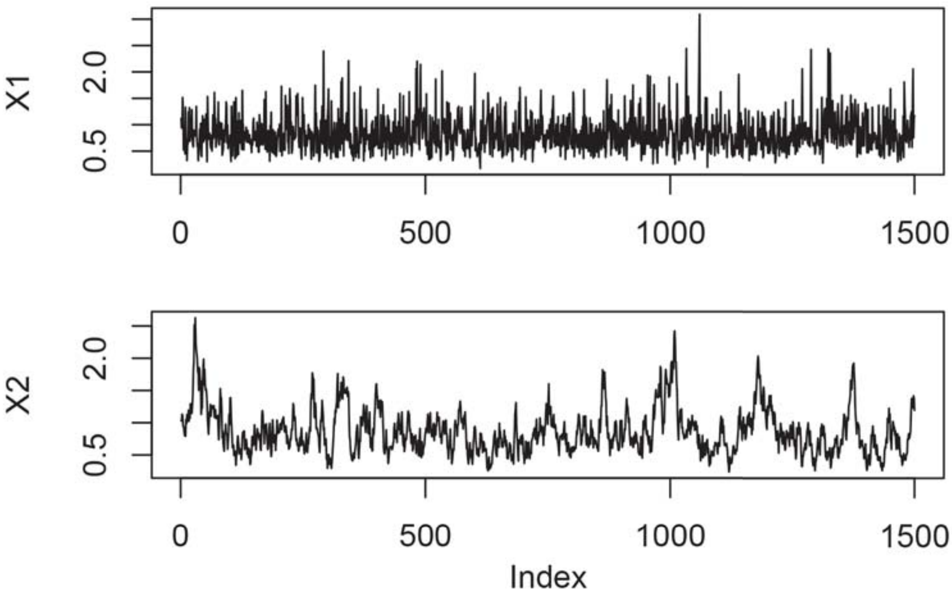

Figure 1: Simulation of $\mathsf{GIG}(1, 2, 3)$-stationary Poisson-driven Markov processes. The top panel displays 1500 simulated data with $\phi = 1$. The bottom repeats the simulation for $\phi = 80$.

for each one of the m values, simulating a realization of the process starting on the last observation and parameterized by such a value. As an example, we computed highest posterior density intervals of probability 0.9 for both series. This appears in Figure 3.

The GEV illustration is in continuous time. We consider two simulated data sets with 1000 observations, the first one with a type I GEV (or Gumbel) distribution with density $f_1(x) = \mathsf{Gum}(x; \mu, \sigma)$, and the second one with a type II GEV distribution with density $f_2(x) = \mathsf{GEV}(x; \mu, \sigma, \nu)$. These cases are obtained via transformations of unitary exponential r.v. and, therefore, the dependency function only involves the parameter c, i.e., $\phi_t = 1/(e^{ct} - 1)$.

We set the parameters $\boldsymbol{\theta}^{(st)} = (\mu, \sigma) = (1, 4)$ for type I GEV, $\boldsymbol{\theta}^{(st)} = (\mu, \sigma, \nu) = (1, 0.8, 0.7)$ for type II GEV, and $\boldsymbol{\theta}^{(tr)} = c = 1$ in both cases. We assign independent unitary exponential priors to σ and ν, and an independent standard normal to μ. The details on the derivation of the relevant posterior distributions appear in the Appendix B.3 (cf. Supporting information).

The KL-divergence between the real invariant distribution and the estimated one is 0.00003 in the I GEV case, and 0.00037 in the II GEV case (see Figure 4). For the c parameter, the posterior distribution mode is given by 1.044 in the I GEV case, and 0.998 in the II GEV case. Therefore, the evidence suggests the estimation method works

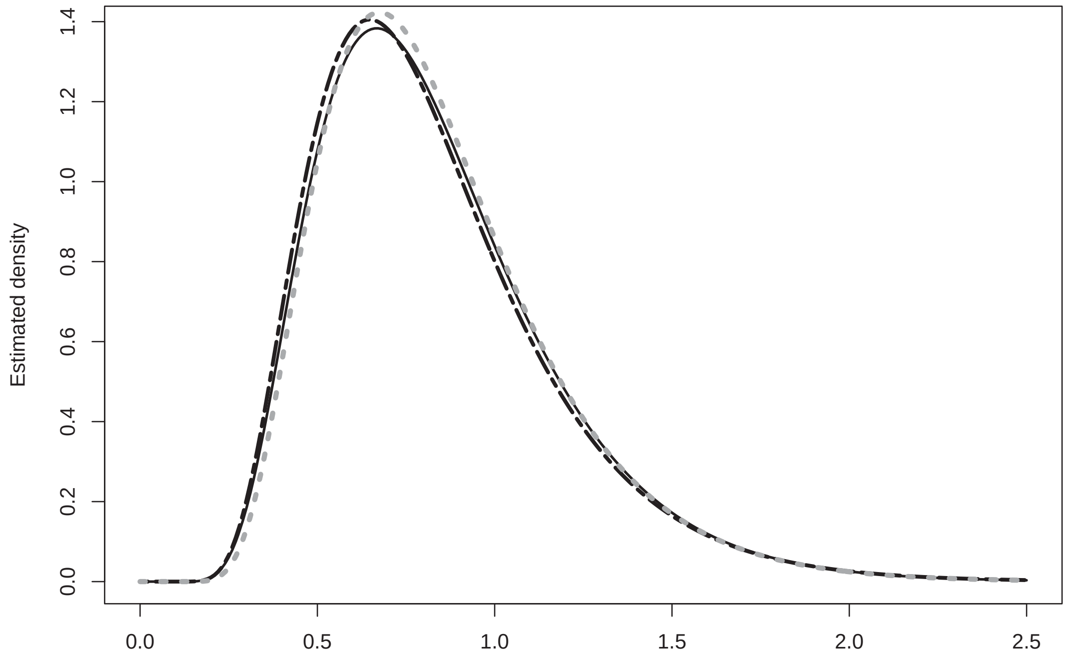

Figure 2: Estimated densities for $\mathsf{GIG}(1,2,3)$-stationary Poisson-driven Markov processes with $\phi = 1$ (black-dashed line) and $\phi = 80$ (grey-dotted line). $\mathsf{GIG}(1,2,3)$ marginal density is also displayed (black-solid line).

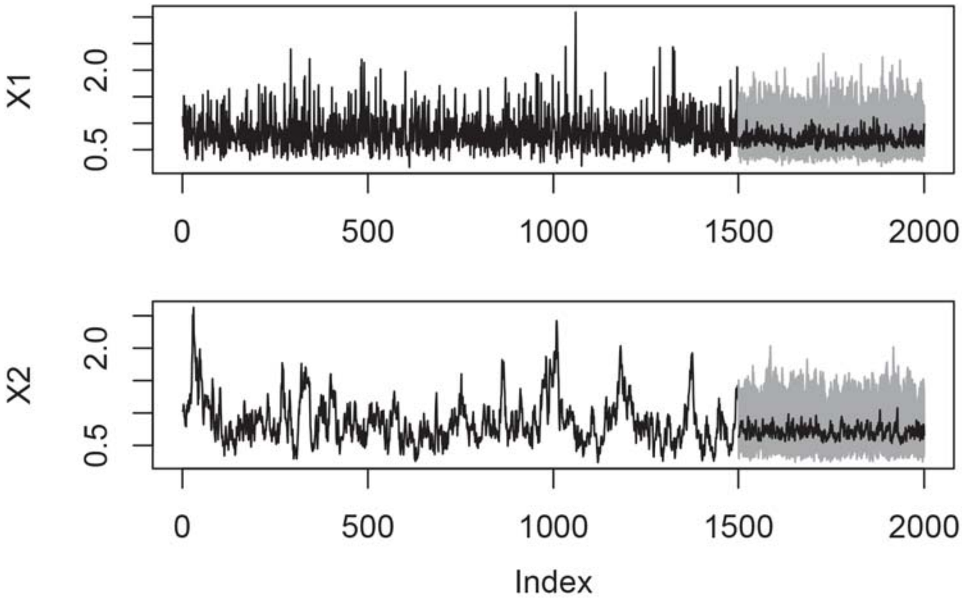

Figure 3: First 1500 observations simulated from a $\mathsf{GIG}(1,2,3)$-stationary Poisson-driven Markov processes for $\phi = 1$ (top panel), and $\phi = 80$ (bottom panel). The posterior predictive mode for 500 observations is also shown (in black) together with its 0.9 probability prediction intervals (shaded-grey). These predictions were computed with 1000 simulated paths.

correctly for type I and II GEV distributions, and we can proceed by applying it to real data.

5.2 Real Data

The models discussed in Sections 2.2 and 3.3 could be an appealing alternative for the econometric analysis of financial series. Indeed, various common stylised features typically observed in these type of data, such as heavy tail distributions and volatility clustering (Cont; 2001), can be well captured with the appropriate choice of stationary density f and through the non-linear dependence driven by ϕ. Here we illustrate how they can capture different features by means of three datasets.

The first dataset consists of 937 daily estimations of the realized volatility of the FTSE 100 equity index, from October 31, 2003 to May 31, 2007. The estimations are those provided in Heber et al. (2009). The open price, with its corresponding log-returns and realized volatility are displayed in Figure 5. Given the positive support, cluster pattern, and heavy tail behavior of the realized volatility, it seems plausible to adopt the discrete-time model with GIG stationary distribution described in Section 2.2.

The data is first cleaned, one outlier is replaced by a missing value, and with it, around four percent of the data is missing seemingly completely at random. The missing data are imputed using the predictive mean matching method provided in Van Buuren and Groothuis-Oudshoorn (2011) R package. After the cleaning, the estimation method is implemented on the realized volatility multiplied by 300. The posterior estimate for the stationary density is displayed in Figure 6. The posterior modes of the model parameters $(\alpha, \delta, \gamma, \phi)$ are equal to $(-2.837, 0.173, 0.316, 37.247)$.

To test the our approach, we break the sample into an estimation period (from October 31, 2003 to December 31, 2006), and a subsequent forecasting period (from January 1, 2007 to May 31, 2007). Next, we predict probability intervals for the forecasting period with 1000 simulated trajectories. The corresponding intervals appear in Figure 7. Indeed, 92 percent of the sample falls within the prediction intervals of probability 0.95. Therefore, we may conclude the method is working correctly for this data.

We delve deeper into the performance of our construction by analyzing two further datasets. We consider the minimum daily stock returns occurring during a month of the S&P 500 and the Tokyo Stock Price Index (TOPIX). The S&P 500 is one of the most representative market indexes and rests upon the common stock prices of 500 top publicly traded American companies. TOPIX measures the market value changes of the common stocks on the Tokyo Stock Exchange. For our analysis the S&P 500 series has a coverage period of almost 12 years, from January 3, 2000 to July 9, 2012, whereas the TOPIX data are based on a 22 years period, from January 1, 1990 to July 31, 2012.

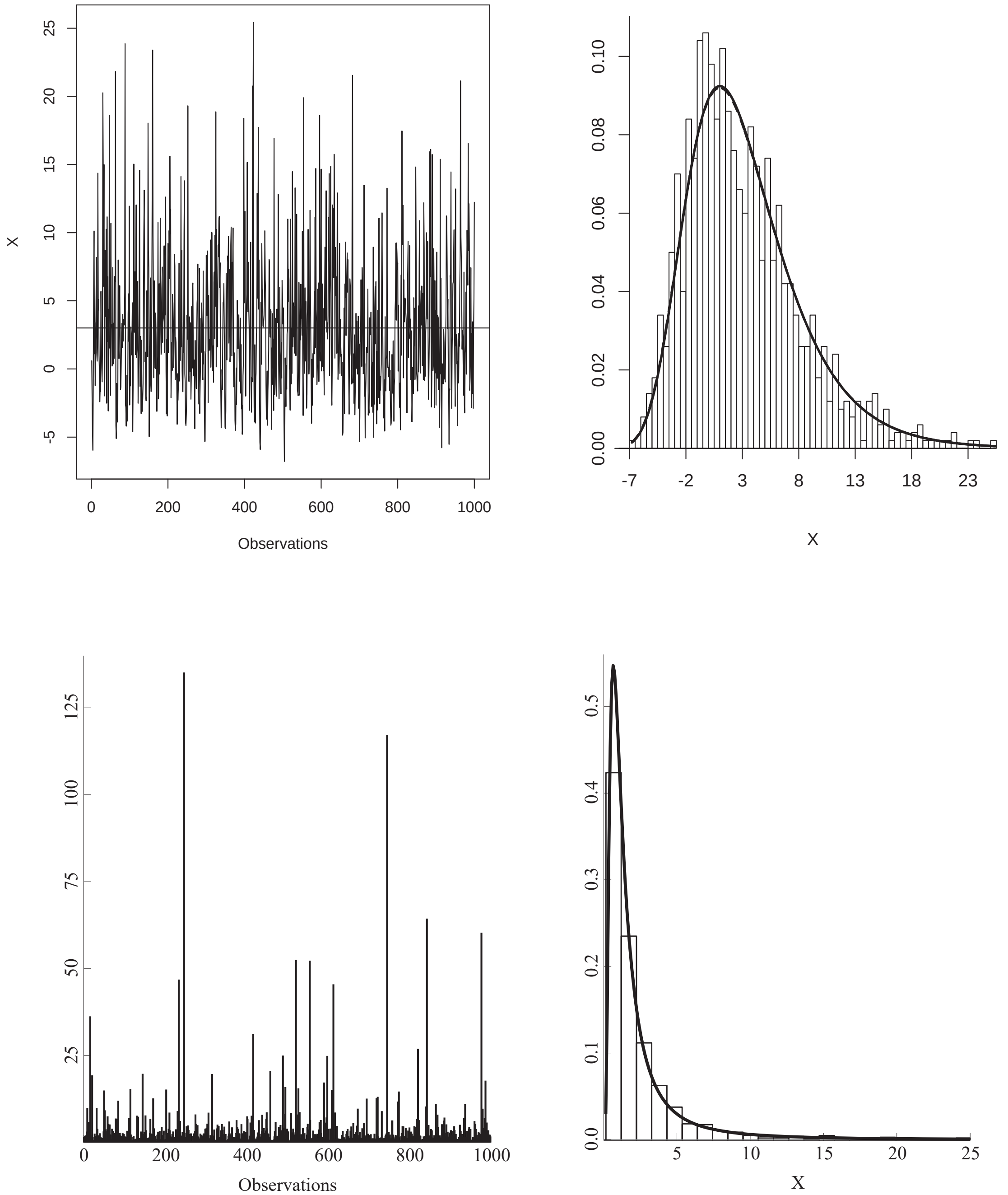

Figure 4: The left panels display 1000 simulated data from type I (top) and type II (bottom) GEV models. The right panels depict the corresponding histograms together with the estimated stationary distributions.

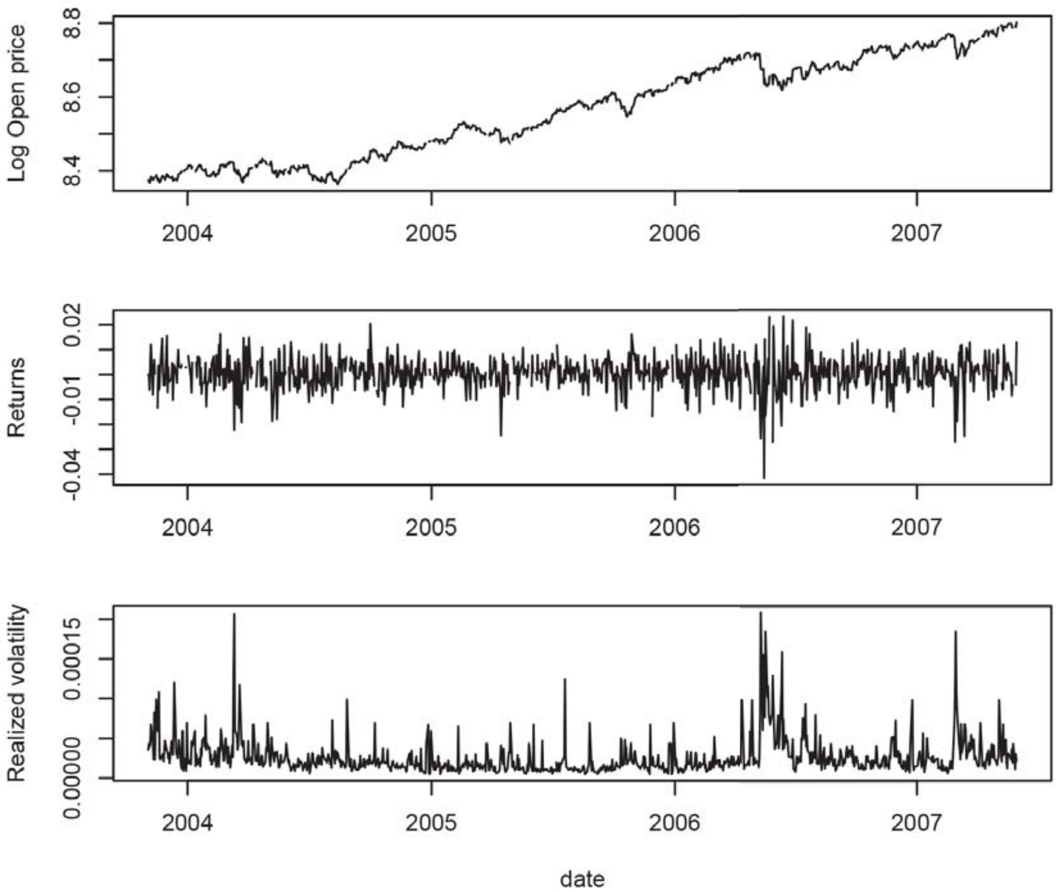

Figure 5: Logarithm of open prices, returns, and realized volatility series for FTSE 100 during the period of October 31, 2003 to May 31, 2007

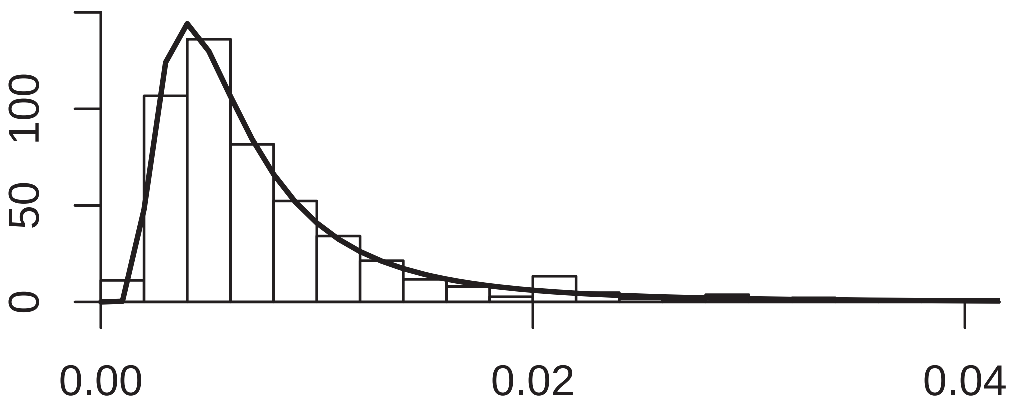

Figure 6: Histogram and density estimate for the stationary distribution based on a GIG–Stationary Poisson driven model for the FTSE 100 data set.

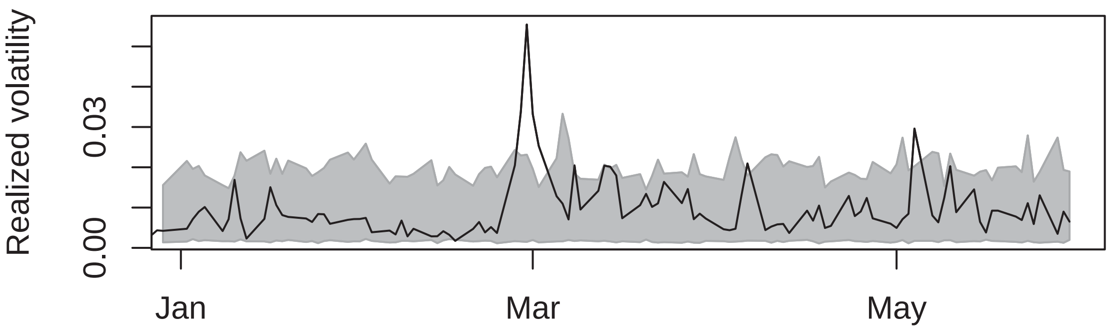

Figure 7: FTSE 100 realized volatility series from January 1, 2007 to May 31, 2007, along with highest posterior density intervals of probability 0.95. The intervals are computed using the estimation period from October 31, 2003 to December 31, 2006, with 1000 simulated trajectories.

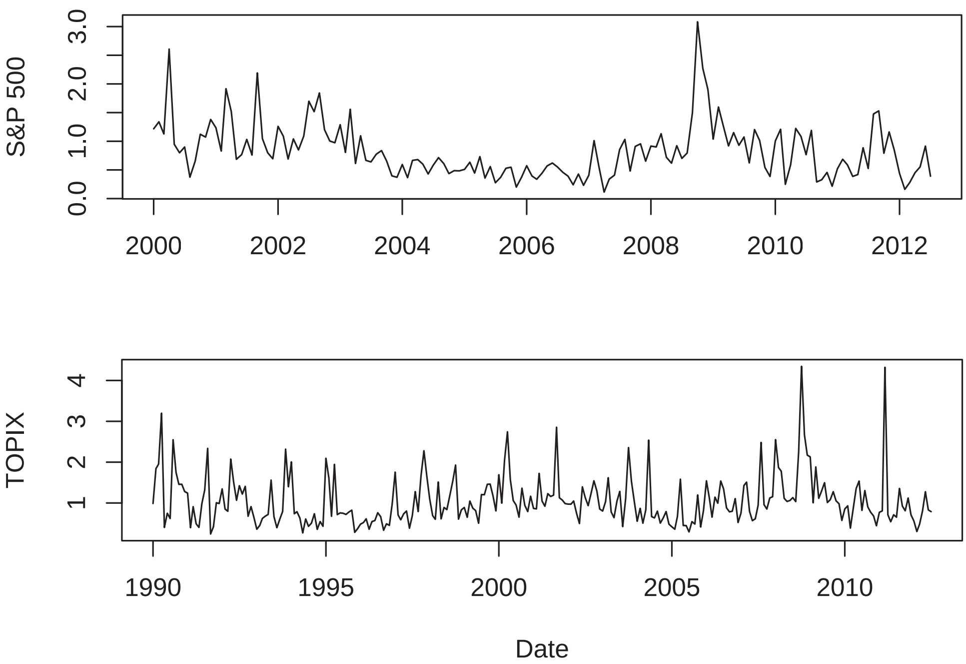

Figure 8: S&P 500 and TOPIX series of log monthly minima from January 3, 2000 to July 9, 2012, and from January 1, 1990 to July 31, 2012, respectively. The values are multiplied by -1 to accommodate the GEV stationary distribution.

In both cases we compute daily returns taking log-differences multiplied by 100 and then compute the monthly minima. Hence, two series consisting of 151 and 271 observations are obtained for S&P 500 and TOPIX, respectively. The extracted series are displayed in Figure 8. Here, it is shown that the heavy tails, clearly observable in both datasets, can be satisfactorily captured by GEV-stationary Poisson-driven Markov processes.

The posterior estimates of the stationary densities are displayed in Figure 9. As to the estimated model parameters, in the type I GEV case the posterior modes of (μ, σ, c) are given by (0.721,0.384,0.828) for S&P 500, and by (0.808, 0.412,0.725) for the TOPIX dataset. In the type II case the posterior modes of the model parameters (μ, σ, ν, c) are (0.618, 0.372, 0.041, 0.399) for the S&P 500 case, and (0.779, 0.385, 0.152, 0.702) for the TOPIX dataset.

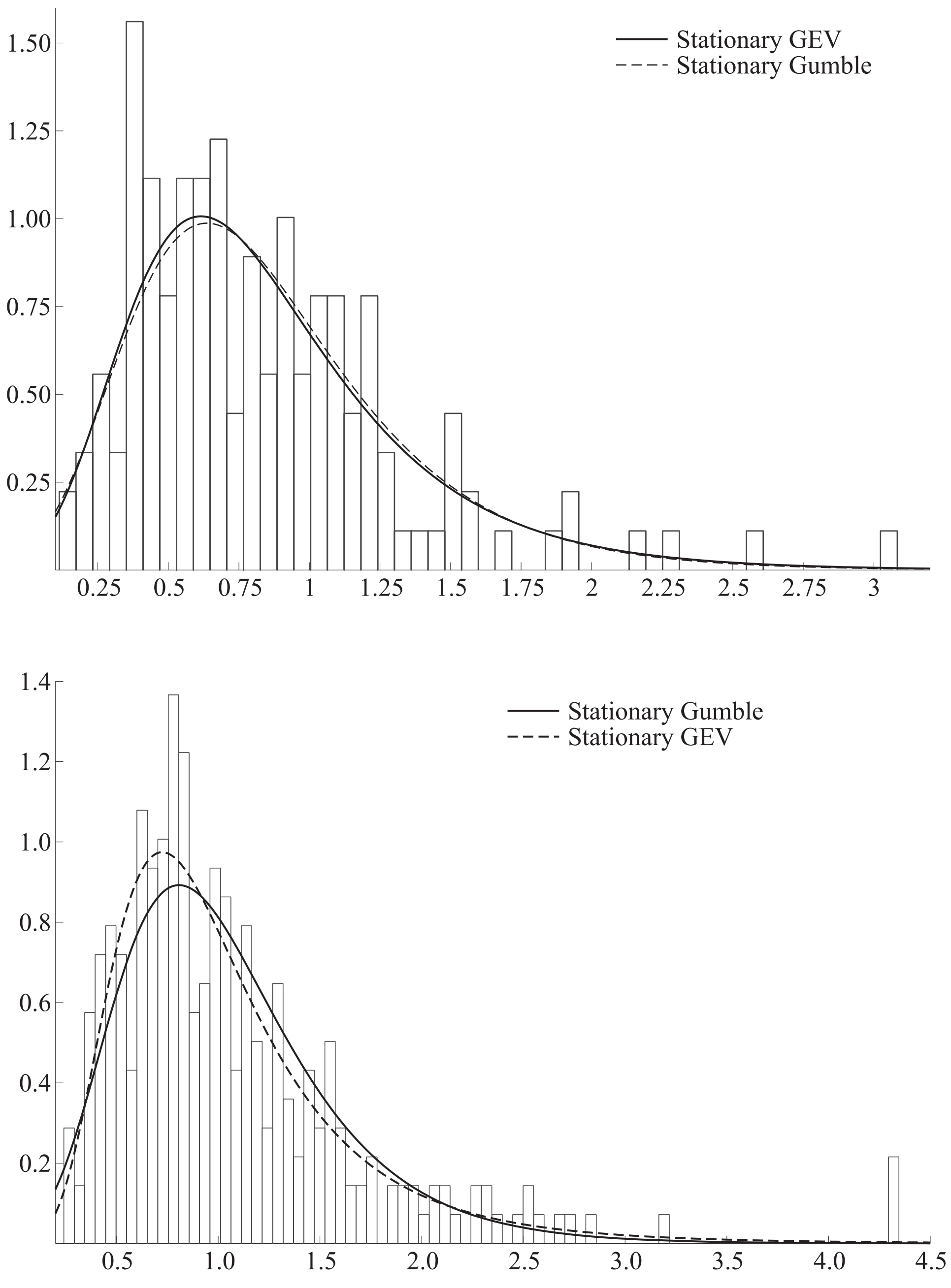

Figure 9: Histograms of the S&P 500 (upper panel) and TOPIX (lower panel) series together with the corresponding estimated stationary densities in the type I (solid line) and type II (dashed line) GEV cases.

6 Concluding remarks

We introduced a novel class of strictly stationary Markov models with arbitrary but given invariant distributions supported on $\mathbb{R}_+$. Given a choice of invariant density f, the dependence in the model is introduced via a Poisson weighed density, which in turn leads to a well defined Markov process. Unlike other existing approaches in the literature, the proposed construction has a useful representation of the underlying transition probability. Although some other constructions, like those in Joe (1996),Jørgensen and Song (1996) and Pitt et al. (2002), allow for great generality, these do not always lead to tractable forms of the transition probability. Indeed, the type of transition mechanism characterizing f-stationary Poisson-driven Markov processes leads to an effective MCMC-based estimation procedure. In particular, this latter aspect is very appealing for the estimation of continuous-time models, where explicit forms of transition probabilities are not always available.

We also showed how the construction can be extended to build new stationary models, whose invariant distributions are supported on other spaces, e.g. taking values on $\mathbb{R}$, without compromising the transition density representation. Particular emphasis was placed on the general classes of GIG and GEV stationary distributions, which themselves constitute interesting choices of models for econometric or financial applications. However, the construction can be applied to any other distribution supported on $\mathbb{R}_+$, not only leading to alternative approaches for parameter-driven or observation-driven stochastic volatility modelling, but also in other areas where model stability is a requirement.

Given that the specific dependence of f-stationary Poisson-driven Markov processes is induced by the choice of invariant density f (besides the fixed contribution already imposed by the Poisson weighed density) one is naturally inclined to choose f as general as possible. Two future research directions we plan to pursue to achieve a high degree of generality for f is to adopt phase-type distributions or nonparametric hierarchical mixtures (Lo; 1984).

Ackowledgements

M. Anzarut and R.H. Mena were partially supported by CONACyT project 241195. I. Prünster is supported by the European Research Council (ERC) through StG “N-BNP” 306406.

Supporting information

Additional information is available online including:

Appendix A Derivation of the results in Section 3.3.

References

Barndorff-Nielsen, O. E. and Shephard, N. (2001). Non-gaussian ornstein-uhlenbeck-based models and some of their uses in financial economics., *J. R. Statist. Soc. B.* **63**: 167–241.

Bibby, M., Skovgaard, M. and Sørensen, M. (2005). Diffusion-type models with given marginal distribution and autocorrelation function., *Bernoulli* **11**: 191–220.

Brockwell, P. J. and Davis, R. A. (1987). *Time Series: Theory and Methods.*, Springer.

Coles, S. (2001). *An Introduction to Statistical Modeling of Extreme Values.*, Springer.

Cont, R. (2001). Empirical properties of asset returns: stylized facts and statistical issues., *Quantitative Finance* **11**: 223–236.

Contreras-Cristán, A., Mena, R. H. and Walker, S. G. (2009). On the construction of stationary ar(1) models via random distributions., *Statistics* **43**: 227–240.

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

Dempster, A., Laird, N. and Rubin, D. (1997). Maximum likelihood from incomplete data via the em algorithm., *J. Roy. Statist. Soc. Ser. B* **39**: 1–38.

Eberlein, E. and Hammerstein, E. A. (2004). Generalized hyperbolic and inverse Gaussian distributions: Limiting cases and approximation of processes., *in* R. Dalang, M. Dozzi and F. Russo (eds), *Seminar on Stochastic Analysis, Random Fields and Applications IV*, Birkhuser Verlag, pp. 221–264.

Embrechts, P., Kluppelberg, C. and Mikosch, T. (1997). *Modelling extremal events for insurance and finance.*, Springer-Verlag.

Ewens, W. (2004). *Mathematical Population Genetics (2nd Edition).*, Springer-Verlag.

Favaro, S., Hadjicharalambous, G. and Prünster, I. (2011). On a class of distributions on the simplex., *J. Statist. Plann. Inference* **141**: 2987–3004.

Favaro, S., Lijoi, A. and Prünster, I. (2012). On the stick-breaking representation of normalized inverse gaussian priors., *Biometrika* **99**: 663–674.

Heber, G., Lunde, A., Shephard, N. and Sheppard, K. K. (2009). *Oxford-Man Institute's realized library, Version 0.2.*, University of Oxford.

Huerta, G. and Sansó, B. (2007). Time-varying models for extreme values., *Environ. Ecol. Stat.* **14**: 285–299.

Joe, H. (1996). Time series models with univariate margins in the convolution-closed infinitely divisible class., *J. Appl. Probab.* **33**: 664–677.

Jørgensen, B. and Song, P. (1996). Stationary time series models with exponential dispersion model margins., *J. Appl. Probab.* **35**: 78–92.

Kalli, M., Griffin, J. E. and Walker, S. G. (2011). Slice sampling mixture models., *Stat. Comput.* **21**: 93–105.

Karlin, S. and Taylor, H. M. (1981). *A Second Course in Stochastic Processes*, Academic Press.

Klüppelberg, C. (2002). Risk management with extreme value theory., *Paper 270*, Institut für Statistick, Ludwig Maximilians Universität, München, Sonderforschungsbereich 386.

Kullback, S. and Leibler, R. A. (1951). On information and sufficiency., *Ann. Math. Statist.* **22**: 79–86.

Linetsky, V. (2007). Spectral methods in derivatives pricing., *in* V. Linetsky (ed.), *Handbook of Financial Engineering, Handbooks in Operations Research and Management Sciences*, Elsevier, Amsterdam.

Lo, A. Y. (1984). On a class of bayesian nonparametric estimates: I. density estimates., *Ann. Statist.* **12**: 351–357.

McKenzie, E. (2003). Discrete variate time series., *in* C. Rao and D. Shanbhag (eds), *Handbook of Statistics.*, Elsevier.

Mena, R. (2001). *Stationary models using latent structures.*, PhD thesis, University of Bath.

Mena, R. H., Ruggiero, M. and Walker, S. G. (2011). Geometric stick-breaking processes for continuous time bayesian nonparametric modeling., *J. Statist. Plann. Inference* **141**: 3217–3230.

Mena, R. H. and Walker, S. G. (2005). Stationary autoregressive models via a bayesian nonparametric approach., *J. Time Series Anal.* **26**: 789–805.

Mena, R. H. and Walker, S. G. (2007a). On the stationary version of the generalized hyperbolic arch model., *Ann. Inst. Statist. Math.* **59**: 325–348.

Mena, R. H. and Walker, S. G. (2007b). Stationary mixture transition distribution (mtd) models via predictive distributions., *J. Statist. Plann. Inference.* **137**: 3103–3112.

Mena, R. H. and Walker, S. G. (2009). On a construction of markov models in continuous time., *Metron* **LXVII**: 303–323.

Nakajima, J., Kunihama, T., Omori, Y. and Frühwirth-Schnatter, S. (2012). Generalized extreme value distribution with time-dependence using the ar and ma models in state space form., *Comput. Statist. Data Anal.* **56**: 3241–3259.

Papaspiliopoulos, O. and Roberts, G. O. (2008). Retrospective mcmc for dirichlet process hierarchical models., *Biometrika.* **95**: 169–186.

Pitt, M. K., Chatfield, C. and Walker, S. G. (2002). Constructing first order autoregressive models via latent processes., *Scand. J. Statist.* **29**: 657–663.

Rao, C. R. (1965). On discrete distributions arising out of methods of ascertainment., *Sankhya* **27**: 311–324.

Van Buuren, S. and Groothuis-Oudshoorn, K. (2011). mice: Multivariate imputation by chained equations in r., *J. of Statistical Software* **45-3**: 1–67.

Yau, C., Papaspiliopoulos, O., Roberts, G. O. and Holmes, C. C. (1965). Bayesian non-parametric hidden markov models with applications in genomics., *J. Roy. Statist. Soc. Ser. B* **73**: 37–57.