Periodic autoregressive stochastic volatility

Aknouche, Abdelhakim

University of Science and Technology Houari Boumediene, Qassim University

1 June 2013

Periodic autoregressive stochastic volatility

Abdelhakim Aknouche *

October 2, 2015

Abstract

This paper proposes a stochastic volatility model ($PAR\text{-}SV$) in which the log-volatility follows a first-order periodic autoregression. This model aims at representing time series with volatility displaying a stochastic periodic dynamic structure, and may then be seen as an alternative to the familiar periodic $GARCH$ process. The probabilistic structure of the proposed $PAR\text{-}SV$ model such as periodic stationarity and autocovariance structure are first studied. Then, parameter estimation is examined through the quasi-maximum likelihood (QML) method where the likelihood is evaluated using the prediction error decomposition approach and Kalman filtering. In addition, a Bayesian $MCMC$ method is also considered, where the posteriors are given from conjugate priors using the Gibbs sampler in which the augmented volatilities are sampled from the Griddy Gibbs technique in a single-move way. As a-by-product, period selection for the $PAR\text{-}SV$ is carried out using the (conditional) Deviance Information Criterion (DIC). A simulation study is undertaken to assess the performances of the QML and Bayesian Griddy Gibbs estimates. Applications of Bayesian $PAR\text{-}SV$ modeling to daily, quarterly and monthly $S\&P$ 500 returns are considered.

Keywords and phrases: Periodic stochastic volatility, periodic autoregression, QML via prediction error decomposition and Kalman filtering, Bayesian Griddy Gibbs sampler, single-move approach, DIC.
Mathematics Subject Classification: AMS 2000 Primary 62M10; Secondary 60F99
Proposed running head: Periodic AR Stochastic volatility.

1. Introduction

Over the past three decades, stochastic volatility (SV) models introduced by Taylor (1982) have played an important role in modelling financial time series which are characterized by a time-varying volatility feature. This class of models is often viewed as a better formal alternative to $ARCH$-type models because the

*Faculty of Mathematics, University of Science and Technology Houari Boumediene, Algiers, Algeria, e-mail: aknouche_ab@yahoo.com.

volatility is itself driven by an exogenous innovation, a fact that is consistent with finance theory, although it makes the model relatively more difficult to estimate. Several extensions of the original SV formulation have been proposed in the literature to account for further volatility features such as long memory, simultaneous dependence, excess kurtosis, leverage effect and change in regime (e.g. Harvey et *al*, 1994; Ghysels et *al*, 1996; Breidt, 1997; Breidt et *al*, 1998; So et *al*, 1998; Chib et *al*, 2002; Carvalho and Lopes, 2007; Omori et *al*, 2007; Nakajima and Omori, 2009). However, it seems that most of the proposed formulations have been devoted to time-invariant volatility parameters and hence they could not meaningfully explain time series whose volatility structure changes over time, in particular volatility displaying a stochastic periodic pattern that cannot be accounted for by time-invariant SV-type models.

In order to describe periodicity in the volatility, Tsiakas (2006) proposed various interesting and parsimonious time-varying stochastic volatility models in which the volatility parameters are expressed as deterministic periodic functions of time with appropriate exogenous variables. The proposed models called "periodic stochastic volatility" (PSV) have been successfully applied to model the evolution of daily $S\&P$ 500 returns. This is an evidence that the periodically changing structure may characterize time series volatility. However, the PSV formulations are by definition especially well adapted to a kind of deterministic periodicity in the second moment and hence they might neglect a possible stochastic periodicity in these moments (see e.g. Ghysels and Osborn, 2001 for the difference between deterministic and stochastic periodicity). A complementary approach which seems to be appropriate in capturing stochastic periodicity in the volatility is to consider a linear time-invariant representation for the volatility equation involving seasonal lags, leading to a seasonal SV specification (see e.g. Ghysels et *al*, 1996). However, because of the time-invariance of the volatility parameters, the seasonal SV model may be too restrictive in representing periodicity and a model with periodic time-varying parameters seems to be more relevant. Indeed, as pointed out by Bollerslev and Ghysels (1996, p. 140) many financial time series encountered in practice are such that neglecting periodic time-variation in the corresponding volatility equation give rise to a loss in forecast efficiency, which is more severe in the $GARCH$ model than in linear $ARMA$. This has motivated Bollerslev and Ghysels (1996) to propose the periodic $GARCH$ (P-$GARCH$) formulation in which the parameters vary periodically over time in order to capture the stochastic periodicity pattern in the conditional second moment. At present the P-$GARCH$ model is among the most important models for describing periodic time series volatility (see e.g. Bollerslev and Ghysels, 1996; Taylor, 2006; Koopman et *al*, 2007; Osborn et *al*, 2008; Regnard and Zakoïan, 2011; Sigauke and Chikobvu, 2011; Aknouche and Al-Eid, 2012). However, despite the recognized relevance of the P-$GARCH$ model, an alternative periodic SV for stochastic periodicity is in fact needed for many reasons. First, it is well known that an SV-like model is more flexible than a $GARCH$ type model because the volatility in the latter is only driven by the past of the observed process which constitutes a restrictive limitation. Second, compared to SV-type models, the probability structure of P-$GARCH$ models is relatively more complex to obtain (Aknouche and Bibi, 2009). Finally, compared to the P-$GARCH$, the PAR-SV easily allows to simple multivariate generalizations.

In this paper we propose to model stochastic periodicity in the volatility through a model that generalizes the standard SV equation so that the parameters vary periodically over time. Thus, in the proposed model termed *periodic autoregressive stochastic volatility* (PAR-SV) the log-volatility process follows a first-order periodic autoregression and may be generalized so as to have any linear periodic representation. This model may be seen as an extension of the models of Tsiakas (2006) to include periodic feature in the autoregressive dynamic of the log-volatility equation. The structure and probability properties of the proposed model such as periodic stationarity, autocovariance structure and relationship with multivariate stochastic volatility models are first studied. In particular, periodic $ARMA$ ($PARMA$) representations for the logarithm of the squared PAR-SV process are proposed. Then, parameter estimation is conducted via the quasi-maximum likelihood (QML) method, properties of which are discussed. In addition, Bayesian estimation approach using Markov Chains Monte Carlo ($MCMC$) techniques is also considered. Specifically, a Gibbs sampler is used to estimate the joint posterior distribution of the parameters and the augmented volatility while calling for the Griddy Gibbs procedure when estimating the conditional posterior distribution of the augmented parameters. On the other hand, selection of the period of the PAR-SV model is carried out using the (conditional) Deviance Information Criterion (DIC). Simulation experiments are undertaken to assess finite-sample performances of the $QMLE$ and the Bayesian Griddy Gibbs methods. Moreover, empirical applications to modeling series of daily, quarterly and monthly $S\&P$ 500 returns are conducted in order to appreciate the usefulness of the proposed PAR-SV model. In the particular daily return case, a variant of the PAR-SV model with missing values, dealing with the "day-of-the-week" effect is applied.

The rest of this paper proceeds as follows. Section 2 proposes the PAR-SV model and studies its main probabilistic properties. In Section 3, the quasi-maximum likelihood method via prediction error decomposition and Kalman filtering is adopted. Moreover, a single-move Bayesian approach by means of the Griddy Gibbs (BGG) sampler is proposed. In particular, some $MCMC$ diagnostic tools are presented and period selection in PAR-SV models is carried out using the DIC. Through a simulation study, Section 4 examines the behavior of the QML and BGG methods in finite samples. Section 5 applies the PAR-SV specification to model daily, quarterly and monthly $S\&P$ 500 returns using the Bayesian Griddy Gibbs method. Finally, Section 6 concludes.

2. The PAR-SV and its main probabilistic properties

In this paper, we say that a stochastic process $\{\varepsilon_t, t \in \mathbb{Z}\}$ has a periodic autoregressive stochastic volatility representation with period S (PAR-SV_S in short) if it is given by

$$\begin{cases} \varepsilon_t = \sqrt{h_t}\eta_t \\ \log(h_t) = \alpha_t + \beta_t \log(h_{t-1}) + \sigma_t e_t \end{cases}, \quad t \in \mathbb{Z}, \tag{2.1a}$$

where the parameters α_t, β_t, and σ_t are S-periodic over t (i.e. $\alpha_t = \alpha_{t+Sn}$ $\forall n \in \mathbb{Z}$ and so on) and the period $S \geq 1$ is the smallest positive integer verifying the latter relationship. $\{(\eta_t, e_t), t \in \mathbb{Z}\}$ is a sequence of independent and identically distributed ($i.i.d.$) random vectors with mean $(0,0)'$ and covariance matrix I_2 (I_2 stands for the identity matrix of dimension 2). We have called model (2.1a) *periodic autoregressive stochastic volatility* rather than shortly *periodic stochastic volatility* because the log-volatility is rather driven by a first-order periodic autoregression and also in order to make distinction between model (2.1a) and the periodic stochastic volatility (PSV) model proposed by Tsiakas (2006). In fact, the PAR-SV model (2.1a) may be generalized so that h_t satisfies any stable periodic $ARMA$ (henceforth $PARMA$) representation.

Note that when $\beta_t = 0$, model (2.1a) reduces to Tsiakas's (2006) model if we take α_t to be an appropriate deterministic periodic function of time. In that case, the effect of any current shock in the innovation e_t influences only the present volatility and does not affect its future evolution. This is the case of what is called *deterministic periodicity.* If, in contrast, $\beta_t \neq 0$ for some t, the log-volatility equation involves lagged values of the log-volatility process. Therefore, the log-volatility consists at any time of an accumulation of past shocks, so that present shocks affect more or less the future log-volatility evolution, depending on the stability of the log-volatility equation (see the periodic stationarity condition (2.5) below). This case is commonly named stochastic periodicity in the volatility.

It should be noted that although h_t is conventionally called volatility, it is not the conditional variance of the observed process given its past information in the familiar sense as in $ARCH$-type models. This is because h_t is instead $\mathcal{F}_t$-measurable and so $E\left(\varepsilon_t^2/\mathcal{F}_{t-1}\right) = E\left(h_t/\mathcal{F}_{t-1}\right) \neq h_t$, where $\mathcal{F}_t$ is the σ-Algebra generated by $\{\varepsilon_u, u \leq t\}$. Nevertheless, $E(h_t) = E\left(\varepsilon_t^2\right)$ and $E\left(\varepsilon_t^2/h_t\right) = h_t$ as in the $ARCH$-type case.

To emphasize the periodicity of the model, let $t = nS + v$ for $n \in \mathbb{Z}$ and $1 \leq v \leq S$. Then model (2.1a) may be written as follows

$$\begin{cases} \varepsilon_{nS+v} = \sqrt{h_{nS+v}}\eta_{nS+v} \\ \log(h_{nS+v}) = \alpha_v + \beta_v \log(h_{nS+v-1}) + \sigma_v e_{nS+v} \end{cases}, n \in \mathbb{Z},\ 1 \leq v \leq S, \tag{2.1b}$$

where by season v ($1 \leq v \leq S$) we mean the channel $\{v, v+S, v+2S, ...\}$ with corresponding parameters α_v, β_v and σ_v.

From (2.1b) the log-volatility appears to be a Markov chain, which is not homogeneous as in time-invariant stochastic volatility models, but is rather periodically homogeneous due to the periodic time-variation of parameters. This may relatively complicate studying the probabilistic structure of the PAR-SV model. As is common in periodic time varying modeling, a routine approach is to write $(2.1b)$ as a time-invariant multivariate SV model by embedding seasons v, $1 \leq v \leq S$ (see e.g. Gladyshev, 1961 and Tiao and Grupe, 1980 for periodic linear models) and then studying the property of this latter. More precisely, define the S-variate sequences $\{H_n, n \in \mathbb{Z}\}$, $\{\boldsymbol{\varepsilon}_n, n \in \mathbb{Z}\}$ by $H_n = (h_{nS+1}, ..., h_{nS+S})'$ and $\boldsymbol{\varepsilon}_n = (\varepsilon_{nS+1}, ..., \varepsilon_{nS+S})'$. Then model $(2.1b)$ may be cast in the following multivariate SV form

$$\begin{cases} \boldsymbol{\varepsilon}_n = diag\left(H_n^{\frac{1}{2}}\right)\Lambda_n \\ \log H_n = B \log H_{n-1} + \xi_n \end{cases}, \quad n \in \mathbb{Z}, \tag{2.2}$$

where $\Lambda_n = \left(\eta_{nS+1}, ..., \eta_{nS+S}\right)'$, $diag\,(a)$ stands for the diagonal matrix formed by the entries of the vector a in the given order. The notations $H_n^{\frac{1}{2}}$ and $\log H_n$ denote the S-vectors defined respectively by $H_n^{\frac{1}{2}}(v) = \sqrt{h_{nS+v}}$ and $\log H_n(v) = \log\left(h_{nS+v}\right)$ $(1 \leq v \leq S)$. The matrices B and ξ_n in (2.2) are given by

$$B = \begin{pmatrix} 0 & \dots & 0 & \beta_1 \\ 0 & \dots & 0 & \beta_2\beta_1 \\ \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & \prod\limits_{v=0}^{S-1}\beta_{S-v} \end{pmatrix}_{S\times S}, \ \xi_n = \begin{pmatrix} \lambda_{1+nS} \\ \beta_2\lambda_{1+nS} + \lambda_{2+nS} \\ \vdots \\ \sum\limits_{k=1}^{S}\prod\limits_{v=0}^{S-k-1}\beta_{S-v}\lambda_{k+nS} \end{pmatrix}_{S\times 1},$$

with $\lambda_{nS+v} = \alpha_v + \sigma_v e_{nS+v}$ $(1 \leq v \leq S)$.

However, this approach has the main drawback that available methods for analyzing multivariate SV models do not consider the particular structure of the coefficients in (2.2) and it may be difficult to conclude on model (2.1). Thus, studying probabilistic and statistical properties of model (2.1) directly may be simpler and better than studying them through model (2.2). This implies that periodic stochastic volatility modelling cannot be trivially deduced from existing multivariate SV analysis. In the sequel, we study the structure of model (2.1) using mainly the direct approach.

Throughout this paper, we frequently use solutions of the following ordinary difference equation

$$u_t = a_t + b_t u_{t-1}, \qquad t \in \mathbb{Z}, \tag{2.3a}$$

with S-periodic coefficients a_t and b_t. Recall that the solution is given, under the requirement that $\left|\prod\limits_{v=0}^{S-1} b_v\right| < 1$, by

$$u_{nS+v} = \left(1 - \prod_{v=0}^{S-1} b_v\right)^{-1} \sum_{j=0}^{S-1}\prod_{i=0}^{j-1} b_{v-i}a_{v-j}, \quad 1 \leq v \leq S,\ n \in \mathbb{Z}. \tag{2.3b}$$

First, we have the following result which provides a necessary and sufficient condition for strict periodic stationarity (see Aknouche and Bibi, 2009 for the definition of strict periodic stationarity).

Theorem 2.1 (Strict periodic stationarity)

The PAR-SV equation given by (2.1) *admits a unique (nonanticipative) strictly periodically stationary and periodically ergodic solution given for $n \in \mathbb{Z}$ and $0 \leq v \leq S-1$ by*

$$\varepsilon_{nS+v} = \eta_{nS+v} \exp\left\{ \frac{1}{2} \left(\frac{\sum_{j=0}^{S-1} \prod_{i=0}^{j-1} \beta_{v-i}\alpha_{v-j}}{1 - \prod_{v=0}^{S-1} \beta_v} + \sum_{j=0}^{\infty} \prod_{i=0}^{j-1} \beta_{v-i}\sigma_{v-j} e_{nS+v-j} \right) \right\}, \tag{2.4}$$

where the series in (2.4) *converges almost surely, if and only if,*

$$\left| \prod_{v=0}^{S-1} \beta_v \right| < 1. \tag{2.5}$$

Proof The result obviously follows from standard linear periodic autoregression (PAR) theory while using (2.3) (see e.g. Aknouche and Bibi, 2009). So, details are omitted. ■

From Theorem 2.1 we see that the monodromy coefficient $\prod_{v=0}^{S-1} \beta_v$ is the analog of the persistent parameter in the case of time-invariant SV and standard $GARCH$ models. If, however, $\left| \prod_{v=0}^{S-1} \beta_v \right| \geq 1$, then clearly there does not exist a nonanticipative strictly periodically stationary solution of (2.1) like (2.4).

Other properties such as periodic geometric ergodicity and strong mixing are obvious. Let first say that a strictly periodically stationary stochastic process $\{\varepsilon_t, t \in \mathbb{Z}\}$ is called *geometrically periodically ergodic* if and only if the corresponding multivariate strictly stationary process $\{\boldsymbol{\varepsilon}_t, t \in \mathbb{Z}\}$ given by $\boldsymbol{\varepsilon}_n = (\varepsilon_{nS+1}, ..., \varepsilon_{nS+S})'$ is geometrically ergodic in the classical sense (see e.g. Meyn and Tweedie, 2009 for the definition of geometric ergodicity).

Theorem 2.2 (Geometric periodic ergodicity)

Under the condition $\left| \prod_{v=0}^{S-1} \beta_v \right| < 1$, the process $\{\varepsilon_t, t \in \mathbb{Z}\}$ defined by (2.1) *is geometrically periodically ergodic. Moreover, if initialized from its invariant measure, then $\{\log h_t, t \in \mathbb{Z}\}$ and hence $\{\varepsilon_t, t \in \mathbb{Z}\}$ are periodically β-mixing with exponential decay.*

Proof The result follows from geometric ergodicity of the vector autoregression $\{\log H_n, n \in \mathbb{Z}\}$ given by (2.2), which may easily be established using Meyn and Tweedie's (2009) results (see also Davis and Mikosch, 2009). ■

Given the form of the periodically stationary solution (2.4), it is easy to give its second-order properties. Assuming the following condition

$$\prod_{j=0}^{\infty} \delta_{v,j} < \infty, \quad \text{for all } 0 \leq v \leq S-1, \tag{2.6}$$

where

$$\delta_{v,j} = E\left(\exp\left(\prod_{i=0}^{j-1} \beta_{v-i}\sigma_{v-j} e_{v-j} \right) \right),$$

we have the following result.

Theorem 2.3 (Second-order periodic stationarity)

Under conditions (2.5) *and* (2.6)*, the series in* (2.4) *converges in the mean square sense and the process given by* (2.4) *is also second-order periodically stationary.*

Proof Routine computation shows that under (2.5) and (2.6) the series in (2.4),

$$\sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{nS+v-j},$$

converges in mean square. Moreover, under these conditions, it is clear that $\{\varepsilon_t, t\in\mathbb{Z}\}$ given by (2.4) is a periodic white noise with periodic variance since $E(\varepsilon_t)=0$, $E(\varepsilon_t\varepsilon_{t-h})=0$ $(h>0)$ and, while using (2.3),

$$\begin{aligned}
Var(\varepsilon_{nS+v}) &= E\left(\exp\left(\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}+\sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{nS+v-j}\right)\right)\\
&= \exp\left(\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right)\prod_{j=0}^{\infty}\delta_{v,j},\quad 0\le v\le S-1. \qquad\blacksquare
\end{aligned}\tag{2.7}$$

In the case of Gaussian log-volatility innovations $\{e_t, t\in\mathbb{Z}\}$, (i.e. $e_t\sim N(0,1)$) it is also possible to obtain more explicit results while reducing assumptions of Theorem 2.3. Using the fact that if $X\sim N(0,1)$ then $E(\exp(\phi X))=\exp(\frac{\phi^2}{2})$ for all non null real constant ϕ, we obtain

$$\delta_{v,j}=\exp\frac{1}{2}\left(\sigma_{v-j}^2\prod_{i=0}^{j-1}\beta_{v-i}^2\right),\tag{2.8}$$

and condition (2.6) of finiteness of $\prod_{j=0}^{\infty}\delta_{v,j}$ reduces to the periodic stationarity condition (2.5): $\left|\prod\limits_{v=0}^{S-1}\beta_v\right|<1$. Moreover, using (2.8) and (2.3) the variance of the process given by (2.7) may be expressed more explicitly as follows

$$\begin{aligned}
Var(\varepsilon_{nS+v}) &= \exp\left(\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right)\prod_{j=0}^{\infty}\exp\left(\frac{1}{2}\sigma_{v-j}^2\prod_{i=0}^{j-1}\beta_{v-i}^2\right)\\
&= \exp\left(\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right)\exp\left(\frac{1}{2}\sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2\right)\\
&= \exp\left(\frac{\sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}+\frac{1}{2}\frac{\sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2}{1-\prod\limits_{v=0}^{S-1}\beta_v^2}\right).
\end{aligned}\tag{2.9}$$

For example, the variance $Var\left(\varepsilon_{nS+v}\right)$ of the process is given respectively, for $S=2$ and $S=3$, by

$$\begin{aligned}
Var\left(\varepsilon_{2n+1}\right) &= \exp\left(\frac{\alpha_1+\beta_1\alpha_2}{1-\beta_1\beta_2}+\frac{1}{2}\frac{\sigma_1^2+\beta_1^2\sigma_2^2}{1-\beta_1^2\beta_2^2}\right),\\
Var\left(\varepsilon_{2n+2}\right) &= \exp\left(\frac{\alpha_2+\beta_2\alpha_1}{1-\beta_1\beta_2}+\frac{1}{2}\frac{\sigma_2^2+\beta_2^2\sigma_1^2}{1-\beta_1^2\beta_2^2}\right),
\end{aligned}$$

$$\begin{aligned}
Var\left(\varepsilon_{3n+1}\right) &= \exp\left(\frac{\alpha_1+\beta_1\alpha_3+\beta_1\beta_3\alpha_2}{1-\beta_1\beta_2\beta_3}+\frac{1}{2}\frac{\sigma_1^2+\beta_1^2\sigma_3^2+\beta_1\beta_3\sigma_2^2}{1-\beta_1\beta_2\beta_3}\right),\\
Var\left(\varepsilon_{3n+2}\right) &= \exp\left(\frac{\alpha_2+\beta_2\alpha_1+\beta_2\beta_1\alpha_3}{1-\beta_1\beta_2\beta_3}+\frac{1}{2}\frac{\sigma_2^2+\beta_2\sigma_1^2+\beta_2\beta_1\sigma_3^2}{1-\beta_1\beta_2\beta_3}\right),\\
Var\left(\varepsilon_{3n+3}\right) &= \exp\left(\frac{\alpha_3+\beta_3\alpha_2+\beta_3\beta_2\alpha_1}{1-\beta_1\beta_2\beta_3}+\frac{1}{2}\frac{\sigma_3^2+\beta_3\sigma_2^2+\beta_3\beta_2\sigma_1^2}{1-\beta_1\beta_2\beta_3}\right).
\end{aligned}$$

Next, the autocovariance of the squared process $\left\{\varepsilon_t^2, t\in\mathbb{Z}\right\}$ is provided. This one is useful in identifying the model and deriving certain estimation methods such as simple and generalized methods of moments. Let $\gamma_v^{\varepsilon^2}(h)=E\left(\varepsilon_{nS+v}^2\varepsilon_{nS+v-h}^2\right)-E\left(\varepsilon_{nS+v}^2\right)E\left(\varepsilon_{nS+v-h}^2\right)$.

Theorem 2.4 (Autocovariance structure of $\left\{\varepsilon_t^2, t\in\mathbb{Z}\right\}$)

i) Under (2.5), (2.6) and the conditions $\prod_{j=0}^{\infty}\delta_{v,j}\delta_{v-h,j}<\infty$ and $E\left(\eta_1^4\right)<\infty$ we have

$$\gamma_v^{\varepsilon^2}(0)=\exp\left(2\frac{\sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right)\left(E\left(\eta_1^4\right)E\left(\exp\left(2\sum_{j=h}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j}\right)\right)-\prod_{j=0}^{\infty}\delta_{v,j}^2\right) \quad (2.10a)$$

$$\begin{aligned}
\gamma_v^{\varepsilon^2}(h)=&\left(E\left(\exp\left(\sum_{j=0}^{h-1}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j}+\left(1+\prod_{i=0}^{h-1}\beta_{v-i}^{-1}\right)\sum_{j=h}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j}\right)\right)-\right.\\
&\left.\prod_{j=0}^{\infty}\delta_{v,j}\delta_{v-h,j}\right)\times\exp\left(\frac{\sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}+\sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right),\ h>0. \quad (2.10b)
\end{aligned}$$

Proof Using (2.4) direct calculation gives

$$\begin{aligned}
E\left(\varepsilon_{nS+v}^2\varepsilon_{nS+v-h}^2\right)=&E\left(\exp\left(\sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{nS+v-j}+\sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\beta_{v-h-i}\sigma_{v-h-j}e_{nS+v-h-j}\right)\right)\\
&\times\exp\left(\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}+\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right)E\left(\eta_{nS+v}^2\eta_{nS+v-h}^2\right), \quad (2.11)
\end{aligned}$$

under finiteness of the latter expectations. When in particular $h=0$, combining (2.7) and (2.11) we get (2.10a) under finiteness of $E\left(\eta_1^4\right)$.

For $h > 0$, because of the independence structure of $\{\eta_t, t \in \mathbb{Z}\}$ one obtains

$$E\left(\varepsilon_{nS+v}^2 \varepsilon_{nS+v-h}^2\right) = \exp\left(\frac{\sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j} + \sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right) \times$$
$$E\left(\exp\left(\sum_{j=0}^{h-1}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j} + \sum_{j=h}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j} + \sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\beta_{v-h-i}\sigma_{v-h-j}e_{v-h-j}\right)\right)$$
$$= \exp\left(\frac{\sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j} + \sum\limits_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\right) \times$$
$$E\left(\exp\left(\sum_{j=0}^{h-1}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j} + \left(1+\prod_{i=0}^{h-1}\beta_{v-i}^{-1}\right)\sum_{j=h}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j}\right)\right),$$

giving (2.10b). ■

Expressions of the S kurtoses $Kurt\,(v)$ $(1 \leq v \leq S)$ of the $PAR\text{-}SV_S$ model may be given from (2.9) and (2.10) by

$$\begin{aligned} Kurt\,(v) &= E\left(\eta_1^4\right)\frac{\prod_{j=0}^{\infty} E\left(\exp\left(\prod\limits_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j}\right)^2\right)}{\prod_{j=0}^{\infty}\left(E\left(\exp\prod\limits_{i=0}^{j-1}\beta_{v-i}\sigma_{v-j}e_{v-j}\right)\right)^2}, \quad 1 \leq v \leq S, \\ &\geq E\left(\eta_1^4\right). \end{aligned} \tag{2.12}$$

By the Cauchy-Schwartz inequality, this clearly shows that the $PAR\text{-}SV$ is characterized by excess Kurtosis for all channels $\{1, ..., S\}$.

In particular, under the normality assumption on the innovations, the second-order periodic stationarity reduces to $E(\eta_1^4) < \infty$ and $\left|\prod\limits_{v=0}^{S-1}\beta_v\right| < 1$. So from (2.8), expression (2.12) reduces to

$$Kurt\,(v) = E\left(\eta_1^4\right), \quad 1 \leq v \leq S.$$

The autocovariance function has also more explicit form in the case of Gaussian $\{e_t, t \in \mathbb{Z}\}$.

Corollary 2.1 (Autocovariance structure of $\left\{\varepsilon_t^2, t \in \mathbb{Z}\right\}$ under normality of $\{e_t, t \in \mathbb{Z}\}$**)**

Under the same assumptions of Theorem 2.4 and if $\{e_t, t \in \mathbb{Z}\}$ *is Gaussian then,*

$$\gamma_v^{\varepsilon^2}(0) = \exp\left(2\left(1-\prod_{v=0}^{S-1}\beta_v\right)^{-1}\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j} + \left(1-\prod_{v=0}^{S-1}\beta_v^2\right)^{-1}\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2\right)\left(E\left(\eta_1^4\right)-1\right), \tag{2.13a}$$

$$\gamma_v^{\varepsilon^2}(h) = \exp\left(\frac{\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}+\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}}{1-\prod_{v=0}^{S-1}\beta_v}+\frac{\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2}{1-\prod_{v=0}^{S-1}\beta_v^2}\right)\times$$
$$\left(\exp\left(\frac{\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2}{1-\prod_{v=0}^{S-1}\beta_v^2}\prod_{i=0}^{h-1}\beta_{v-i}\right)-1\right), \qquad h>0. \qquad (2.13b)$$

Proof For Gaussian innovations, we use again the fact that if $X\sim N(0,1)$ then $E(\exp(\phi X))=\exp(\frac{\phi^2}{2})$. Therefore, (2.13$a$) follows from (2.10$a$) and (2.9). For $h>0$ we have

$$E\left(\varepsilon_{nS+v}^2\varepsilon_{nS+v-h}^2\right)=\exp\left(\left(1-\prod_{v=0}^{S-1}\beta_v\right)^{-1}\left(\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}+\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}\right)\right)\times$$
$$\prod_{j=0}^{h-1}\exp\left(\frac{1}{2}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2\right)\prod_{j=0}^{\infty}\exp\left(\frac{1}{2}\left(1+\prod_{i=0}^{h-1}\beta_{v-i}^{-1}\right)^2\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2\right).$$

After tedious but straightforward calculation, the autocovariance function at lag h ($h>0$) simplifies for Gaussian innovations to

$$\gamma_v^{\varepsilon^2}(h) = \exp\left(\frac{\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}+\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}}{1-\prod_{v=0}^{S-1}\beta_v}\right)\left[\exp\left(\sum_{j=0}^{h-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\frac{\sigma_{v-j}^2}{2}\right)\times\right.$$
$$\left.\exp\left(\sum_{j=h}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2\left(1+\prod_{i=0}^{h-1}\beta_{v-i}^{-1}\right)^2\right)-\exp\left(\sum_{j=0}^{\infty}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2\right)\right].$$
$$= \exp\left(\frac{\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j}+\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-h-i}\alpha_{v-h-j}}{1-\prod_{v=0}^{S-1}\beta_v}+\frac{\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2}{1-\prod_{v=0}^{S-1}\beta_v^2}\right)\times$$
$$\left(\exp\left(\frac{\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2}{1-\prod_{v=0}^{S-1}\beta_v^2}\prod_{i=0}^{h-1}\beta_{v-i}\right)-1\right),$$

which is (2.13b). ■

It is worth noting that expanding the exponential function in (2.13b) under the periodic stationarity condition (2.5), the autocovariance function $\gamma_v^{\varepsilon^2}(h)$ of the squared process $\left\{\varepsilon_t^2, t\in\mathbb{Z}\right\}$ has the following equivalent form as $h\to\infty$

$$\gamma_v^{\varepsilon^2}(h)\sim K\prod_{i=0}^{h-1}\beta_{v-i}\sim K\left(\prod_{v=0}^{S-1}\beta_v\right)^{h/S},$$

and so $\gamma_v^{\varepsilon^2}(h)$ converges geometrically to zero as $h\to\infty$, where K is an appropriate real constant. However, the decreasing of $\gamma_v^{\varepsilon^2}(h)$ is not compatible with the recurrence equation that satisfy periodic $ARMA$

$(PARMA)$ autocovariances and we can conclude that the squared process $\left\{\varepsilon_t^2, t \in \mathbb{Z}\right\}$ does not admit a $PARMA$ autocovariance representation.

Nevertheless, the logarithmed squared process $\left\{\log\left(\varepsilon_t^2\right), t \in \mathbb{Z}\right\}$ has in fact a $PARMA$ autocovariance structure. Considering the following notations $Y_t = \log\left(\varepsilon_t^2\right)$, $X_t = \log h_t$, $u_t = \log\left(\eta_t^2\right)$, $E\left(\log\left(\eta_t^2\right)\right) = \mu_u$ and $Var\left(\log\left(\eta_t^2\right)\right) = \Delta_u^2$, we have from (2.1)

$$Y_t = X_t + u_t. \tag{2.14}$$

Theorem 2.5 ($PARMA(1,1)$ representation of $\left\{\log\left(\varepsilon_t^2\right), t \in \mathbb{Z}\right\}$)

Under assumption (2.5) *and finiteness of* Δ_u^2, *the process* $\{Y_t, t \in \mathbb{Z}\}$ *has a* $PARMA_S(1,1)$ *representation given by*

$$Y_{nS+v} - \mu_v^Y = \beta_v\left(Y_{nS+v-1} - \mu_{v-1}^Y\right) + \zeta_{nS+v} - \psi_v \zeta_{nS+v-1}, \qquad 1 \le v \le S,\, t \in \mathbb{Z}, \tag{2.15a}$$

where $\mu_v^Y = E\left(Y_{nS+v}\right)$,

$$\psi_v = \begin{cases} \frac{\left(1+\beta_v^2\right)\Delta_u^2 + \sigma_v^2 - \sqrt{\left(\left(1+\beta_v^2\right)\Delta_u^2+\sigma_v^2\right)\left(\left(1-\beta_v^2\right)\Delta_u^2+\sigma_v^2\right)}}{2\beta_v\Delta_u^2} & \text{if } \Delta_u^2 \neq 0 \\ 0 & \text{if } \Delta_u^2 = 0 \end{cases}, \ 1 \le v \le S, \tag{2.15b}$$

and $\{\zeta_t, t \in \mathbb{Z}\}$ *is a periodic white noise with periodic variance*

$$\sigma_{\zeta,v}^2 = Var\left(\zeta_{nS+v}\right) = \begin{cases} \frac{\beta_v\Delta_u^2}{\psi_v} & \text{if } \prod_{v=0}^{S-1}\beta_v \neq 0 \\ 0 & \text{if } \prod_{v=0}^{S-1}\beta_v = 0 \end{cases}, \ 1 \le v \le S. \tag{2.15c}$$

Proof The second-order structure of $\{X_t, t \in \mathbb{Z}\}$ is given form (2.1) while using (2.3),

$$\mu_v^X = E\left(X_{nS+v}\right) = \alpha_v + E\left(X_{nS+v-1}\right) = \left(1 - \prod_{v=0}^{S-1}\beta_v\right)^{-1}\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j},$$

$$\gamma_v^X(0) = Var\left(X_{nS+v}^2\right) = \beta_v^2 E\left(X_{nS+v-1}^2\right) + \sigma_v^2 = \left(1 - \prod_{v=0}^{S-1}\beta_v^2\right)^{-1}\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2$$

$$\begin{aligned}\gamma_v^X(h) &= Cov\left(X_{nS+v}, X_{nS+v-h}\right) \\ &= \beta_v\gamma_{v-1}^X(h-1),\end{aligned} \quad , \ h > 0.$$

Therefore, using (2.14) we have

$$\mu_v^Y = E\left(Y_{nS+v}\right) = E\left(X_{nS+v}\right) + E\left(u_{nS+v}\right) = \left(1 - \prod_{v=0}^{S-1}\beta_v\right)^{-1}\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}\alpha_{v-j} + \mu_u,$$

$$\gamma_v^Y(0) = Var\left(Y_{nS+v}\right) = Var\left(X_{nS+v}\right) + \Delta_u^2 = \left(1 - \prod_{v=0}^{S-1}\beta_v^2\right)^{-1}\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-i}^2\sigma_{v-j}^2 + \Delta_u^2$$

$$\begin{aligned}&\gamma_v^Y(h) = \gamma_v^X(h) = \beta_v\gamma_{v-1}^X(h-1) = \beta_v\beta_{v-1}...\beta_{v-h+1}\gamma_{v-h}^X(0) \\ &= \beta_v\beta_{v-1}...\beta_{v-h+1}\left(1 - \prod_{v=0}^{S-1}\beta_v^2\right)^{-1}\sum_{j=0}^{S-1}\prod_{i=0}^{j-1}\beta_{v-h-i}^2\sigma_{v-h-j}^2\end{aligned} \quad , \ h > 0.$$

Clearly, the process $\{Y_t, t \in \mathbb{Z}\}$ has a $PARMA$ representation since $\gamma_v^Y(h) = \beta_v \gamma_{v-1}^Y(h-1)$, for $h > 1$. To identify the parameters of its representation we use expressions of $\gamma_v^Y(h)$ for $h = 0, 1$. If $\{Y_t, t \in \mathbb{Z}\}$ has a $PARMA$ representation (2.15a) then for all $1 \leq v \leq S$

$$\begin{aligned} \gamma_v^Y(0) &= \beta_v \gamma_v^Y(1) + \sigma_{\zeta,v}^2 \left(1 + \psi_v \left(\psi_v - \beta_v\right)\right) \\ \gamma_v^Y(1) &= \beta_v \gamma_{v-1}^Y(0) - \psi_v \sigma_{\zeta,v}^2. \end{aligned} \tag{2.15d}$$

Hence, if $\Delta_u^2 \neq 0$ we have for all $1 \leq v \leq S$,

$$\begin{aligned} \frac{1 + \psi_v \left(\psi_v - \beta_v\right)}{\psi_v} &= \frac{\gamma_v^Y(0) - \beta_v \gamma_v^Y(1)}{\beta_v \gamma_{v-1}^Y(0) - \gamma_v^Y(1)} \\ &= \frac{\gamma_v^Y(0) - \beta_v^2 \gamma_{v-1}^X(0)}{\beta_v \left(\gamma_{v-1}^X(0) + \Delta_u^2\right)\right) - \beta_v \gamma_{v-1}^X(0)} \\ &= \frac{\gamma_v^Y(0) - \left(\gamma_v^X(0) - \sigma_v^2\right)}{\beta_v \Delta_u^2} \\ &= \frac{\gamma_v^Y(0) - \left(\gamma_v^Y(0) - \sigma_v^2 - \Delta_u^2\right)}{\beta_v \Delta_u^2} \\ &= \frac{\sigma_v^2 + \Delta_u^2}{\beta_v \Delta_u^2}. \end{aligned} \tag{2.15e}$$

The latter equation admits, for all $1 \leq v \leq S$, two solutions one of which is with modulus less than 1 ($|\psi_v| < 1$) and is given by (2.15b). Such a choice clearly ensures that $\prod_{v=0}^{S-1} |\psi_v| < 1$, but it is not unique. Moreover, when $\prod_{v=0}^{S-1} \beta_v \neq 0$ using (2.15d), the variance of $\{\zeta_t, t \in \mathbb{Z}\}$ is

$$\begin{aligned} \sigma_{\zeta,v}^2 &= \frac{\beta_v \gamma_{v-1}^Y(0) - \gamma_v^Y(1)}{\psi_v} \\ &= \frac{\beta_v \left(\gamma_{v-1}^X(0) + \Delta_u^2\right) - \beta_v \gamma_{v-1}^Y(0)}{\psi_v}, \end{aligned}$$

showing (2.15c).

If, however, $\Delta_u^2 = 0$ the relationship $\gamma_v^Y(h) = \beta_v \gamma_{v-1}^Y(h-1)$ also holds for $h = 1$ and so the process $\{Y_t, t \in \mathbb{Z}\}$ is a pure first-order periodic autoregression ($PAR(1)$) with $\psi_v = 0$ for all v. When $\prod_{v=0}^{S-1} \beta_v = 0$, the process $\{Y_t, t \in \mathbb{Z}\}$ is a strong periodic white noise (an independent and periodically distributed, *i.p.d.* sequence) and so $\psi_v = 0$ for all v (see also Francq and Zakoïan, 2006 for the particular non-periodic case $S = 1$). ■

It is worth noting that representation (2.15a) is not unique. Indeed, in contrast with time-invariant $ARMA$ models for which an $ARMA$ process may be uniquely identified from its autocovariance function (see Brockwell and Davis, 1991), it is not always possible to build a unique $PARMA$ model from an autocovariance function having $PARMA$ structure. However, we may enumerate all possible representations from solving (2.15d) and choosing the best one fitting the observed series. The resulting representation will be abusively said the $PARMA$ representation. Such a representation is useful for obtaining predictions for the process $\left\{\log\left(\varepsilon_t^2\right), t \in \mathbb{Z}\right\}$. It may also be used to obtain approximate predictions for the squared process $\left\{\varepsilon_t^2, t \in \mathbb{Z}\right\}$ as this latter does not admit a $PARMA$ representation (see Section 4.2). If we denote by $\widehat{\varepsilon}_{t+h/t}^2 = E\left(\varepsilon_{t+h}^2/\varepsilon_t^2, \varepsilon_{t-1}^2, ...\right)$ the mean-square prediction of ε_{t+h}^2 based on $\varepsilon_t^2, \varepsilon_{t-1}^2, ...$, then $\widehat{\varepsilon}_{t+h/t}^2$ may be approximated by

$$C \exp\left(\widehat{\log\left(\varepsilon_{t+h/t}^2\right)}\right),$$

where

$$\widehat{\log\left(\varepsilon_{t+h/t}^2\right)} = E\left(\log\left(\varepsilon_{t+h}^2\right) / \log\left(\varepsilon_t^2\right), \log\left(\varepsilon_{t-1}^2\right), ...\right),$$

and C is a normalization factor. The constant C is introduced to minimize the bias due to using incorrectly the following relationship

$$\exp\left(\widehat{\log\left(\varepsilon_{t+h/t}^2\right)}\right) = \exp\widehat{\log\left(\varepsilon_{t+h/t}^2\right)},$$

as we know from Jensen's inequality that the latter equality is in fact not true. Typically, one can take C as the sample variance of $\left(\log\left(\varepsilon_t^2\right), t = 1, ..., T\right)$.

3. Parameter estimation of the PAR-SV model

In this Section we consider two estimation methods for the PAR-SV model. The first one is a QML method based on prediction-error decomposition of a corresponding linear periodic state-space model. This method which uses Kalman filtering to obtain linear predictors and error prediction variances is used as a Benchmark to the second proposed method, which is based on the Bayesian approach. In this method, from given conjugate priors, the conditional posteriors are obtained from the Gibbs sampler in which the conditional posterior of the augmented volatilities is driven via the Griddy-Gibbs technique. In the rest of this Section we consider a series $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ generated from model (2.1) with sample-size $T = NS$ supposed without loss of generality multiple of the period S. The vector of model parameters is denoted by $\theta = \left(\omega', \sigma^{2\prime}\right)'$ where $\omega = \left(\omega_1', \omega_2', ..., \omega_S'\right)'$, $\omega_v = \left(\alpha_v, \beta_v\right)'$ and $\sigma^2 = \left(\sigma_1^2, \sigma_2^2, ..., \sigma_S^2\right)'$.

3.1 $QMLE$ via prediction error decomposition and Kalman filtering

Taking in (2.1) the logarithm of the square of ε_t we obtain the following linear periodic state space-model

$$\left\{\begin{array}{l} Y_{nS+v} = \mu + X_{nS+v} + \widetilde{u}_{nS+v} \\ X_{nS+v} = \alpha_v + \beta_v X_{nS+v-1} + \sigma_v e_{nS+v} \end{array}\right., \quad n \in \mathbb{Z},\ 1 \le v \le S, \tag{3.1}$$

where as in the above $Y_{v+nS} = \log\left(\varepsilon_{nS+v}^2\right)$, $X_{nS+v} = \log\left(h_{nS+v}\right)$, $u_{nS+v} = \log\left(\eta_{nS+v}^2\right)$, $\mu = E\left(u_{nS+v}\right)$, $\widetilde{u}_{nS+v} = u_{nS+v} - \mu$ and $\Delta_u^2 = Var\left(u_{nS+v}\right)$. When $\{\eta_t, t \in \mathbb{Z}\}$ is standard Gaussian, the mean and variance of $\log\left(\eta_{nS+v}^2\right)$ can accurately be approximated by $\psi\left(\frac{1}{2}\right)-\ln\left(\frac{1}{2}\right)\approx -1.27$ and $\pi^2/2$ respectively, where $\psi\left(.\right)$ is the gamma function. Note, however, that the linear state-space model (3.1) is not Gaussian, unless i) e_1 is Gaussian, ii) e_1 and η_1 are independent and iii) η_1 has the same distribution as $\exp\left(X/2\right)$ for some X normally distributed with mean zero and variance 1. In what follows we assume for simplicity of exposition that η_1 is standard Gaussian, but the QML method we present below is still valid when η_1 is not Gaussian and even when μ and Δ_u^2 are unknown.

Let $Y=(Y_1,\ldots,Y_T)'$ be the series of log-squares corresponding to $\varepsilon=(\varepsilon_1,\ldots,\varepsilon_T)'$ (i.e. $Y_t=\log\left(\varepsilon_t^2\right)$, $1\le t\le T$), which is generated from (3.1) with true parameter θ_0. The quasi-likelihood function $l_Q(\theta;Y)$ evaluated at a generic parameter θ may be written via the prediction error decomposition as follows

$$\log(l_Q(\theta;Y))=-\frac{T}{2}\log(2\pi)-\frac{1}{2}\sum_{t=1}^{T}\left(\log(F_t)+\frac{(Y_t-\widehat{Y}_{t/t-1})^2}{F_t}\right), \tag{3.2}$$

where $\widehat{Y}_{t/t-1}=\widehat{X}_{t|t-1}+\mu$, $\widehat{X}_{t|t-1}$ is the best predictor of the state X_t based on the observations $Y_1,...,Y_{t-1}$ with mean square errors $P_{t/t-1}=E\left(X_t-\widehat{X}_{t/t-1}\right)^2$ and $F_t=E\left(Y_t-\widehat{Y}_{t/t-1}\right)^2$. A QML estimate $\widehat{\theta}_{QML}$ of the true θ_0 is the maximizer of $\log(l_Q(\theta;Y))$ over some compact parametric space Θ, where $l_Q(\theta;Y)$ it is evaluated as if the linear state space model (3.1) was Gaussian. Thus the best state predictor $\widehat{X}_{t|t-1}$ and the state prediction error variance $P_{t/t-1}$ may be recursively computed using the Kalman filter, which in the context of model (3.1) is described by the following recursions

$$\left.\begin{array}{l}\widehat{X}_{t/t-1}=\beta_t\left(\widehat{X}_{t-1/t-2}+P_{t-1/t-2}F_{t-1}^{-1}\left(Y_{t-1}-\widehat{X}_{t-1/t-2}-\mu\right)\right)+\alpha_t\\ P_{t/t-1}=\beta_t^2\left(P_{t-1/t-2}-P_{t-1/t-2}^2F_{t-1}^{-1}\right)+\sigma_t^2\\ F_t=P_{t/t-1}+\Delta_u^2\end{array}\right.,\quad 2\le t\le T, \tag{3.3a}$$

while remembering that α_t, β_t and σ_t^2 are S-periodic over t. The start-up values of (3.3a) are calculated on the basis of: $\widehat{X}_{1/0}=E\left(X_1\right)$ and $P_{1/0}=Var\left(X_1\right)$. Using results of Section 2, we then get

$$\widehat{X}_{1/0}=\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{1-i}\alpha_{1-j}}{1-\prod\limits_{v=0}^{S-1}\beta_v}\text{ and }P_{1/0}=\frac{\sum_{j=0}^{S-1}\prod\limits_{i=0}^{j-1}\beta_{1-i}^2\sigma_{1-j}^2}{1-\prod\limits_{v=0}^{S-1}\beta_v^2}. \tag{3.3b}$$

Recursions (3.3) may also be used in a reverse form for smoothing purposes, i.e. to obtain the best linear predictor $\widetilde{X}_t$ of X_t based on $Y_1,\ldots,Y_T$, from which we get estimates of the unobserved volatilities h_t $(1\le t\le T)$.

Consistency and asymptotic normality of the QML estimate may be established using standard theory of linear (non-Gaussian) signal plus noise models with time-invariant parameters (Dunsmuir, 1979). For this, we invoke the corresponding multivariate time-invariant model (2.2) which we transform to a linear form as follows

$$\left\{\begin{array}{l}\mathbf{Y}_n=\log H_n+\Xi_n\\ \log H_n=B\log H_{n-1}+\xi_n\end{array}\right.,\quad n\in\mathbb{Z}, \tag{3.4}$$

where $\mathbf{Y}_n$ and Ξ_n are S-vectors such that $\mathbf{Y}_n(v) = Y_{v+nS}$, and $\Xi_n(v) = u_{v+nS}$ $(1 \leq v \leq S)$ and where $\log H_n$, B and ξ_n are given by (2.2). Using (3.4), we can call for the theory in Dunsmuir (1979) to yield the asymptotic variance of the $QMLE$ under the finiteness of the moment $E\left(Y_{v+nS}^4\right)$ (see also Ruiz, 1994 and Harvey et *al*, 1994).

Of course, the $QMLE$ would be asymptotically efficent if we assume that e_t is Gaussian, e_1 and and η_1 are independent, and has η_1 the same distribution as $\exp(X/2)$, where $X \sim N(0,1)$. In that case, $\log\left(\eta_1^2\right) \sim N(0,1)$ and the lienar state space (3.1) would be also Gaussian. Therefore, the $QMLE$ reduces to the exact maximum likelihood estimate (MLE). However, the assumption that $\log\left(\eta_1^2\right) \sim N(0,1)$ seems to have a little interest in practice.

3.2. Bayesian inference via Gibbs sampling

Adopting the Bayesian approach, the parameter vector θ of the model and the unobserved volatilities $h = (h_1, h_2, ..., h_T)'$ which are also considered as augmented parameters, are viewed as random with a certain prior distribution $f(\theta, h)$. Given a series $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_T)'$ generated from the $PAR\text{-}SV_S$ model (2.1) with Gaussian innovations, the goal is to make inference about the joint posterior distribution, $f(\theta, h/\varepsilon)$, of (θ, h) given ε. Because of the periodic structure of the $PAR\text{-}SV$ model it is natural to assume that the parameters h, ω, $\sigma_1^2, \sigma_2^2, ..., \sigma_S^2$ are independent of each other. Thus, the joint posterior distribution $f(\theta, h/\varepsilon) = f\left(\omega, \sigma^2, h/\varepsilon\right)$ can be estimated using Gibbs sampling provided we can draw samples from any of the $S+2$ conditional posterior distributions $f\left(\omega/\varepsilon, \sigma^2, h\right)$, $f\left(\sigma_v^2/\varepsilon, \omega, \sigma_{-\{v\}}^2, h\right)$ $(1 \leq v \leq S)$ and $f\left(h/\varepsilon, \omega, \sigma^2\right)$, where $x_{-\{t\}}$ denotes the vector obtained from x after removing its t-th component x_t. Since the posterior distribution of the volatility parameter $f\left(h/\varepsilon, \omega, \sigma^2\right)$ has a rather complicated expression, we sample it element-by-element as done by Jacquier et *al* (1994). Thus, the Gibbs sampler for sampling from the joint posterior distribution $f\left(\omega, \sigma^2, h/\varepsilon\right)$ reduces to drawing samples from any of the $T+S+1$ conditional posterior distributions $f\left(\omega/\varepsilon, \sigma^2, h\right)$, $f\left(\sigma_v^2/\varepsilon, \omega, \sigma_{-\{v\}}^2, h\right)$, $(1 \leq v \leq S)$ and $f\left(h_t/\varepsilon, \omega, \sigma^2, h_{-\{t\}}\right)$, $(1 \leq t \leq T)$. Under normality of the volatility proxies and using standard linear regression theory with an appropriate adaptation to the PAR form of the log-volatility equation (2.1), the conditional posteriors $f\left(\omega/\varepsilon, \sigma^2, h\right)$ and $f\left(\sigma_v^2/\varepsilon, \omega, \sigma_{-\{v\}}^2, h\right)$, $(1 \leq v \leq S)$ may be determined directly from given conjugate priors $f(\omega)$ and $f\left(\sigma_v^2\right)$, $(1 \leq v \leq S)$. However, like the non-periodic SV case (Jacquier et *al*, 1994), direct draws from the distribution $f\left(h_t/\varepsilon, \omega, \sigma^2, h_{-\{t\}}\right)$ are not possible because it has unusual form. Nevertheless, unlike Jacquier et *al* (1994) which used a Metropolis-Hasting chain after determining the form of $f\left(h_t/\varepsilon, \omega, \sigma^2, h_{-\{t\}}\right)$ except for a scaling factor, we use the Griddy-Gibbs procedure as in Tsay (2010) because in our periodic context its implementation seems much simpler.

3.2.1. Prior and posterior sampling analysis

a) Sampling the log-volatility periodic autoregressive parameter ω Before giving the conditional posterior distribution $f\left(\omega/\varepsilon,\sigma^2,h\right)$ through some conjugate prior distributions and linear regression theory, we first write the PAR log-volatility equation as a standard linear regression. Setting $\mathcal{H}_{nS+v} = \left(\underbrace{0,...,0}_{v-1 \text{ times}}, 1, \log\left(h_{nS+v-1}\right), \underbrace{0,...,0}_{S-v \text{ times}}\right)'$, model $(2.1b)$ for $t=1,...,NS$ may be rewritten in the following periodically homoskedastic linear regression

$$\log\left(h_{nS+v}\right) = \mathcal{H}'_{nS+v}\omega + \sigma_v e_{nS+v}, \quad 1 \leq v \leq S,\ 0 \leq n \leq N-1, \tag{3.5a}$$

or also as a standard regression

$$\frac{\log\left(h_{nS+v}\right)}{\sigma_v} = \frac{1}{\sigma_v}\mathcal{H}'_{nS+v}\omega + e_{nS+v}, \quad 1 \leq v \leq S,\ 0 \leq n \leq N-1, \tag{3.5b}$$

with *i.i.d.* Gaussian errors. Assuming known the variances σ_v^2 $(1 \leq v \leq S)$ and the initial observation h_0, the least squares estimate $\widehat{\omega}_{WLS}$ of ω, based on $(3.5b)$, (which is just the weighted least squares estimate of ω based on $(3.5a)$) has the following form

$$\widehat{\omega}_{WLS} = \left(\sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{1}{\sigma_v^2}\mathcal{H}_{nS+v}\mathcal{H}'_{nS+v}\right)^{-1}\sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{1}{\sigma_v^2}\mathcal{H}_{nS+v}\log\left(h_{nS+v}\right),$$

and is normally distributed with mean ω and covariance matrix

$$\Gamma = \left(\sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{1}{\sigma_v^2}\mathcal{H}_{nS+v}\mathcal{H}'_{nS+v}\right)^{-1}. \tag{3.6}$$

Under assumption $(3.5b)$, information of the data about ω is contained in the weighted least squares estimate $\widehat{\omega}_{WLS}$ of ω. To get a closed-form expression for the conditional posterior $f\left(\omega/\varepsilon,\sigma^2,h\right)$ we use a conjugate prior for ω. This prior distribution is Gaussian, i.e. $\omega \sim N\left(\omega^0,\Sigma^0\right)$, where the hyperparameters ω^0,Σ^0 are known and are fixed so that to have a quite reasonably diffuse prior yet informative.

Thus, using standard regression theory (Box and Tiao, 1973; Tsay, 2010) the conditional posterior distribution of ω given ε,σ^2,h is:

$$\omega/\varepsilon,\sigma^2,h \sim N\left(\omega^*,\Sigma^*\right), \tag{3.7a}$$

where

$$\Sigma^* = \left(\sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{1}{\sigma_v^2}\mathcal{H}_{nS+v}\mathcal{H}'_{nS+v} + \left(\Sigma^0\right)^{-1}\right)^{-1} \tag{3.7b}$$

$$\omega^* = \Sigma^*\left(\sum_{n=0}^{N-1}\sum_{v=1}^{S}\frac{1}{\sigma_v^2}\mathcal{H}_{nS+v}\log\left(h_{nS+v}\right) + \left(\Sigma^0\right)^{-1}\omega^0\right). \tag{3.7c}$$

Some remarks are in order:

i) The matrix Γ given by (3.6) is block diagonal. So if we assume that Σ^0 is also block diagonal, then we obtain the same result as if we assume that the seasonal parameters $\omega_1, \omega_2, ..., \omega_S$ are independent of each other, and each one has a conjugate prior with hyperparameters, say ω_v^0 and Σ_v^0 $(1 \leq v \leq S)$, that are appropriate components of ω^0 and Σ^0.

ii) Faster and more stable computation of ω^* and Σ^* in (3.7) which does not involve any matrix inversion (in contrast with (3.7b)) may be obtained while setting $\omega^* = \omega_{NS}^*$, $\Sigma^* = \Sigma_{NS}^*$ and recursively then computing the latter quantities using the well-known recursive least squares (RLS) algorithm (see Ljung and Söderström, 1983, Lemma 2.2) which is given by

$$\begin{array}{l} \omega_{nS+v}^* = \omega_{nS+v-1}^* + \dfrac{\Sigma_{nS+v-1}^{*-1}\mathcal{H}_{nS+v}\left(\log\left(h_{nS+v}\right) - \mathcal{H}_{nS+v}'\omega_{nS+v-1}^*\right)}{\sigma_v^2 + \mathcal{H}_{nS+v}'\Sigma_{nS+v-1}^{*-1}\mathcal{H}_{nS+v}} \\ \Sigma_{nS+v}^{*-1} = \Sigma_{nS+v-1}^{*-1} - \dfrac{\Sigma_{nS+v-1}^{*-1}\mathcal{H}_{nS+v}\mathcal{H}_{nS+v}'\Sigma_{nS+v-1}^{*-1}}{\sigma_v^2 + \mathcal{H}_{nS+v}'\Sigma_{nS+v-1}^{*-1}\mathcal{H}_{nS+v}} \end{array}, \quad \begin{array}{c} 1 \leq v \leq S \\ 0 \leq n \leq N-1, \end{array} \qquad (3.8a)$$

with starting values

$$\omega_0^* = \omega_0 \quad \text{and} \quad \Sigma_0^{*-1} = \Sigma_0. \qquad (3.8b)$$

This may improve the numerical stability and computation time tied to the whole estimation method, especially for a large period S.

b) Sampling the log-volatility periodic variance parameters σ_v^2, $1 \leq v \leq S$ We also use conjugate priors for σ_v^2, $1 \leq v \leq S$ to get a closed form expression for the conditional posterior of σ_v^2 given data and the other parameters $\sigma_{-\{v\}}^2$. Such priors are provided by the inverted *Khi*-squared distribution:

$$\tfrac{a_v\lambda_v}{\sigma_v^2} \sim \chi_{a_v}^2, \qquad 1 \leq v \leq S, \qquad (3.9a)$$

where $a_v\lambda_v = 1$ $(1 \leq v \leq S)$. Given the parameters ω and h, if we define

$$e_{nS+v} = \log\left(h_{nS+v}\right) - \alpha_v - \beta_v \log\left(h_{nS+v-1}\right), \qquad 1 \leq v \leq S,\ 0 \leq n \leq N-1, \qquad (3.9b)$$

then $e_v, e_{v+S}, ..., e_{(N-1)S+v} \sim iiN\left(0, \sigma_v^2\right)$, $1 \leq v \leq S$. From standard Bayesian linear regression theory (see e.g. Tsay, 2010) the conditional posterior distribution of σ_v^2, $1 \leq v \leq S$, given the data and the remainder parameters is an inverted *Khi*-squared distribution with degree of freedom $a_v + N - 1$, that is

$$\tfrac{a_v\lambda_v + \sum_{n=0}^{N-1} e_{nS+v}^2}{\sigma_v^2} / \varepsilon, \omega, \sigma_{-\{v\}}^2, h \sim \chi_{a_v+N-1}^2, \qquad 1 \leq v \leq S. \qquad (3.9c)$$

c) Sampling the augmented volatility parameters $\underline{h} = (h_1, h_2, ..., h_T)'$ Now, it remains to sample from the conditional posterior distribution $f\left(h_t/\varepsilon, \theta, h_{-\{t\}}\right)$, $t = 1, 2, ..., T$. Let us first give the expression of this distribution (except for a multiplicative constant) and we will show how to (indirectly) draw samples from it using the Griddy Gibbs technique. Because of the Markovian (but non-homogeneous) structure of the volatility process $\{h_t, t \in \mathbb{Z}\}$ and the conditional independence of ε_t and h_{t-h} $(h \neq 0)$ given h_t, it follows that for any $1 < t < T$.

$$\begin{aligned} f\left(h_t/\varepsilon, \theta, h_{-\{t\}}\right) &= \frac{f\left(h_t/h_{t-1}, \theta\right) f\left(h_{t+1}/h_t, \theta\right) f\left(\varepsilon_t/\theta, h_t\right)}{f\left(h_{t+1}/h_{t-1}, \theta\right) f\left(\varepsilon_t/\theta, h_{t-1}, h_{t+1}\right)} \\ &\propto f\left(h_t/h_{t-1}, \theta\right) f\left(h_{t+1}/h_t, \theta\right) f\left(\varepsilon_t/\theta, h_t\right). \end{aligned} \tag{3.10}$$

Using the fact that $\varepsilon_t/\theta, h_t \equiv \varepsilon_t/h_t \sim N(0, h_t)$, $\log(h_t)/\log(h_{t-1}), \theta \sim N\left(\alpha_t + \beta_t \log(h_{t-1}), \sigma_t^2\right)$, and $d\log(h_t) = \frac{1}{h_t} dh_t$, formula (3.10) becomes

$$f\left(h_t/\varepsilon, \theta, h_{-\{t\}}\right) \propto \frac{1}{\sqrt{h_t^3}} \exp\left(-\frac{\varepsilon_t^2}{2h_t} - \frac{1}{2\Omega_t}\left(\log(h_t) - \mu_t\right)^2\right), \quad 1 < t < T, \tag{3.11a}$$

where

$$\mu_t = \frac{\sigma_{t+1}^2\left(\alpha_t + \beta_t \log(h_{t-1})\right) + \sigma_t^2 \beta_{t+1}\left(\log(h_{t+1}) - \alpha_{t+1}\right)}{\sigma_{t+1}^2 + \sigma_t^2 \beta_{t+1}^2} \tag{3.11b}$$

$$\Omega_t = \frac{\sigma_{t+1}^2 \sigma_t^2}{\sigma_{t+1}^2 + \sigma_t^2 \beta_{t+1}^2}. \tag{3.11c}$$

Note that in (3.11a) we have used the well-known formula (see Box and Tiao, 1973, p. 418) $A(x-a)^2 + B(x-b)^2 = (x-c)^2(A+B) + (a-b)^2 \dfrac{AB}{A+B}$, where $c = (Aa + Bb)/(A+B)$ provided that $A + B \neq 0$.

For the two end-points h_1 and h_T we may simply use a naive approach which consists of assuming h_1 fixed so that the sampling starts with $t = 2$ and use the fact that $\log(h_T)/\theta, \log(h_{T-1}) \sim N(\alpha_T + \beta_{T-1}\log(h_{T-1}), \sigma_T^2)$. Alternatively, we may also use a forecast of h_{T+1} and a backward prediction of h_0 and employ again formula (3.11) for $0 < t < T+1$. In that case, we forecast h_{T+1}, on the basis of the log-volatility equation of model (2.1), by using a 2-step ahead forecast $\widehat{\log(h_{T-1})}(2)$, at the origin $T-1$, which is given from (2.1) by $\widehat{\log(h_{T-1})}(2) = \alpha_{T+1} + \beta_{T+1}\alpha_T + \beta_{T+1}\beta_T \log(h_{T-1})$. The backward forecast of h_0 is obtained using a 2-step ahead backward forecast on the basis of the backward periodic autoregression (Sakai and Ohno, 1997) associated to the PAR log-volatility.

Once the conditional posterior $f\left(h_t/\varepsilon, \theta, h_{-\{t\}}\right)$ is determined except for a scale factor, we may use some indirect sampling algorithms to draw the volatility h_t. Jacquier et *al* (1994) used the rejection Metropolis-Hasting algorithm. Alternatively, following Tsay (2010) we call for the Griddy-Gibbs technique (Ritter and Tanner, 1992) which consists in:

i) Choosing a grid of m points from a given interval $[h_{t1}, h_{tm}]$ of h_t: $h_{t1} \leq h_{t2} \leq ... \leq h_{tm}$; then evaluating the conditional posterior $f\left(h_t/\varepsilon, \theta, h_{-\{t\}}\right)$ via (3.11) (ignoring the normalization constant) at each one of these points, giving $f_{ti} = f\left(h_{ti}/\varepsilon, \theta, h_{-\{t\}}\right)$, $i = 1, ..., m$.

ii) Building from the values $f_{t1}, f_{t2}, ..., f_{tm}$ the discrete distribution $p(.)$ defined at h_{ti} $(1 \leq i \leq m)$ by $p(h_{ti}) = \dfrac{f_{ti}}{\sum_{j=1}^m f_{tj}}$. This may be seen as an approximation to the inverse cumulative distribution of $f\left(h_t/\varepsilon, \theta, h_{-\{t\}}\right)$.

iii) Generating a number from the uniform distribution on $(0,1)$ and transforming it using the discrete distribution $p(.)$ obtained in ii) to get a random draw for h_t.

It is worth noting that the choice of the grid $[h_{t1}, h_{tm}]$ is crucial for efficiency of the Griddy algorithm. We follow here a similar device by Tsay (2010), which consists of taking the range of h_t, at the l-th Gibbs iteration, to be $[h_t^{*m}, h_t^{*M}]$, where

$$h_t^{*m} = 0.6 \max\left(h_t^{(0)}, h_t^{(l-1)}\right), \;\; h_t^{*M} = 1.4 \min\left(h_t^{(0)}, h_t^{(l-1)}\right), \tag{3.12}$$

$h_t^{(l-1)}$ and $h_t^{(0)}$ being, respectively, the estimate of h_t for the $(l-1)$-th iteration and initial value.

3.2.2. Bayes Griddy Gibbs sampler for PAR-SV

The following algorithm summarizes the Gibbs sampler for drawing from the conditional posterior distribution $f(\theta, h/\varepsilon)$ given ε. For $l = 0, 1, ..., M$, consider the notation $h^{(l)} = \left(h_1^{(l)}, ..., h_T^{(l)}\right)'$, $\omega^{(l)} = \left(\alpha_1^{(l)}, \beta_1^{(l)}, ..., \alpha_S^{(l)}, \beta_S^{(l)}\right)'$ and $\sigma^{2(l)} = \left(\sigma_1^{2(l)}, \sigma_2^{2(l)}, ..., \sigma_S^{2(l)}\right)'$.

Algorithm 3.1

Step 0 Specify starting values $h^{(0)}$, $\omega^{(0)}$ and $\sigma^{2(0)}$.

Step 1 Repeat for $l = 0, 1, ..., M-1$,

Draw $\omega^{(l+1)}$ from $f\left(\omega/\varepsilon, \sigma^{2(l)}, h^{(l)}\right)$ using $(3.7a)$ and (3.8).

Draw $\sigma^{2(l+1)}$ from $f\left(\sigma^2/\varepsilon, \omega^{(l+1)}, h^{(l)}\right)$ using $(3.9b)$ and $(3.9c)$.

Repeat for $t = 1, 2, ..., T = NS$

Griddy Gibbs:

Select a grid of m points $\left(h_{ti}^{(l+1)}\right)$: $h_{t1}^{(l+1)} \leq h_{t2}^{(l+1)} \leq ... \leq h_{tm}^{(l+1)}$.

For $1 \leq i \leq m$ calculate $f_{ti}^{(l+1)} = f\left(h_{ti}^{(l+1)}/\varepsilon, \theta^{(l)}, h_{-\{t\}}^{(l)}\right)$ from (3.11).

Define the inverse distribution $p\left(h_{ti}^{(l+1)}\right) = \dfrac{f_{ti}^{(l+1)}}{\sum_{j=1}^{m} f_{tj}^{(l+1)}}$, $1 \leq i \leq m$.

Generate a number u from the uniform $(0,1)$ distribution.

Transform u using the inverse distribution $p(.)$ to get $h_t^{(l+1)}$, which is considered as a draw from $f\left(h_t/\varepsilon, \theta^{(l+1)}, h_{-\{t\}}^{(l)}\right)$.

Step 2 Return the values $h^{(l)}$, $\omega^{(l)}$ and $\sigma^{2(l)}$, $l = 1, ..., M$. ■

3.2.3. Inference and prediction using the Gibbs sampler for PAR-SV

Once sampling from the posterior distribution $f(\theta, h/\varepsilon)$, statistical inference for the PAR-SV model may be easily made.

The Bayes Griddy-Gibbs parameter estimate $\widehat{\theta}_{BGG}$ of θ is taken to be the posterior mean $\overline{\theta} = E(\theta/\varepsilon)$ which is, under the Markov chain ergodic theorem, approximated with any desired degree of accuracy by

$$\widehat{\theta}_{BGG} = \frac{1}{M} \sum_{l=l_0}^{M+l_0} \theta^{(l)},$$

where $\theta^{(l)}$ is the l-th draw of θ from $f(\theta, h/\varepsilon)$ given by Algorithm 3.1, l_0 is the burn-in size, i.e. the number of initial draws discarded, and M is the number of draws.

Smoothing and forecasting volatility are obtained as a by-product of the Bayes Griddy-Gibbs method. The smoothed value, $\overline{h}_t = E(h_t/\varepsilon)$, of h_t $(1 \leq t \leq T)$ is obtained while sampling from the distribution $f(h_t/\varepsilon)$ which in turn is the marginal of the posterior distribution $f(\theta, h/\varepsilon)$. So $E(h_t/\varepsilon)$ may be accurately approximated by $\frac{1}{M}\sum_{l=l_0}^{M+l_0} h_t^{(l)}$ where $h_t^{(l)}$ is the l-th draw of h_t from $f(\theta, h_t/\varepsilon)$. Forecasting future values $h_{T+1}, h_{T+2}, .., h_{T+k}$ are getting either as in the above using the log-volatility equation with the Bayes parameter estimates or directly while sampling from the predictive distribution $f(h_{T+1}, h_{T+2}, .., h_{T+k}/\varepsilon)$ (see also Jacquier et *al*, 1994).

3.2.4 $MCMC$ diagnostics

It is important to discuss the numerical properties of the proposed BGG method in which the volatilities are sampled element by element. Despite the ease of implementation, it is well documented that the main drawback of the single-move approach (e.g. Kim et *al*, 1998) is that the posterior draws are often highly correlated thereby resulting in a slow mixing and so a slow convergence properties. Among several $MCMC$ diagnostic measures, we consider here the Relative Numerical Inefficiency (RNI) (e.g. Geweke, 1989; Geyer, 1992), which is given by

$$RNI = 1 + 2\sum_{k=1}^{B} K\left(\tfrac{k}{B}\right)\widehat{\rho}_k,$$

where $B = 500$ is the bandwidth, $K(.)$ is the Parzen kernel (e.g. Kim et *al*, 1998) and $\widehat{\rho}_k$ the sample autocorrelation at lag k of the BGG parameter draws. The RNI indicates in fact on the inefficiency due to the serial correlation of the BGG draws (see also Geweke, 1989; Tsiakas, 2006). Another $MCMC$ diagnostic measure (Geweke, 1989) we use here is the Numerical Standard Error (NSE), which is the square root of the estimated asymptotic variance of the $MCMC$ estimator. In fact, the NSE is given by

$$NSE = \sqrt{\tfrac{1}{M}\left(\widehat{\gamma}_0 + 2\sum_{k=1}^{B} K\left(\tfrac{k}{B}\right)\widehat{\gamma}_k\right)},$$

where $\widehat{\gamma}_k$ is the sample autocovariance at lag k of the BGG parameter draws and M is the number of draws.

3.2.5 Period selection via the Deviance Information Criterion

An important issue in PAR-SV modeling is the selection of the period S. This problem is especially more pronounced for modeling daily returns because their periodicity is not as obvious as in quarterly or monthly data. Although many authors (e.g. Franses and Paap, 2000; Tsiakas, 2006) have emphasized the day-of-the-week effect in daily stock returns, which often entails a period of $S = 5$, the period selection problem in periodic volatility models remains a challenging problem. Standard order selection measures such as the AIC and BIC, which require the specification of the number of free parameters in each model, are not applicable for comparing complex Bayesian hierarchical models like the PAR-SV model. This is because in the PAR-SV model, the number of free parameters, which is augmented by the latent volatilities that are in fact not independent but Markovian, is not well defined (cf. Berg et *al*, 2004). For a long time, the Bayes factor has been viewed as the best way to carry out Bayesian model comparison. However, its calculation based on evaluating the marginal likelihood requires extremely high-dimensional integration, and this would be more computationally demanding especially for PAR-SV model which involves a larger number of parameters augmented by the volatilities, exceeding the sample size.

In this paper, we will carry out period selection using rather the *Deviance Information Criterion* (DIC), which may be viewed as a trade-off between model adequacy and model complexity (Spiegelhalter et *al*, 2002). Such a criterion, which represents a Bayesian generalization of the AIC, is easily obtained from $MCMC$ draws, needing no extra-calculations. The (conditional) DIC as introduced by Spiegelhalter et *al* (2002) is defined in the context of PAR-SV_S to be

$$DIC\left(S\right) = -4E_{\theta,h/\varepsilon}\left(\log\left(f\left(\varepsilon/\theta,h\right)\right)\right) + 2\log\left(f\left(\varepsilon/\overline{\theta},\overline{h}\right)\right),$$

where $f\left(\varepsilon/\theta,h\right)$ is the (conditional) likelihood of the PAR-SV model for a given period S and $\left(\overline{\theta},\overline{h}\right) = E\left((\theta,h)/\varepsilon\right)$ is the posterior mean of (θ,h). From the Griddy-Gibbs draws, the expectation $E_{\theta,h/\varepsilon}\left(\log\left(f\left(\varepsilon/\theta,h\right)\right)\right)$ can be estimated by averaging the conditional log-likelihood, $\log f\left(\varepsilon/\theta,h\right)$, over the posterior draws of (θ,h). Further, the joint posterior mean estimate of $(\overline{\theta},\overline{h})$ can be approximated by the mean of the posterior draws of $(\theta^{(l)},h^{(l)})$. Using the fact that $f\left(\varepsilon/\theta,h\right) := f\left(\varepsilon/h\right) = -\frac{1}{2}\sum_{t=1}^{T}\left(\log\left(2\pi h_t\right) + \frac{\varepsilon_t^2}{h_t}\right)$, the $DIC\left(S\right)$ is estimated by

$$\frac{2}{M}\sum_{l=l_0}^{l_0+M}\sum_{t=1}^{T}\left(\log\left(2\pi h_t^{(l)}\right) + \frac{\varepsilon_t^2}{h_t^{(l)}}\right) - \sum_{t=1}^{T}\left(\log\left(2\pi\overline{h}_t\right) + \frac{\varepsilon_t^2}{\overline{h}_t}\right),$$

where $h_t^{(l)}$ denotes the l-th BGG draw of h_t from $f\left(h_t/\varepsilon_t,\theta\right)$, M is the number of draws, l_0 is the burn-in size and $\overline{h}_t := E\left(h_t/\varepsilon\right)$ is estimated by $\frac{1}{M}\sum_{l=l_0}^{l_0+M} h_t^{(l)}$ $(1 \leq t \leq n)$. Of course, a model is preferred if it has the smallest DIC value.

Since the DIC is random and for the same fitted series it may change value from a $MCMC$ draw to another, it is useful to get its corresponding numerical standard error. However, as pointed out by Berg et *al*

(2004), non efficient method has been developed for calculating reasonably accurate Monte Carlo standard errors of DIC. Nevertheless, following the recommendation of Zhu and Carlin (2000) we simply replicate the calculation of DIC some G times and estimate $Var(DIC)$ by its sample variance, giving a broad indication of the implied variability of DIC.

Note finally that for the class of latent variable models to which belongs the $PAR\text{-}SV$, there are in fact several alternative definitions of the DIC depending on the different concepts of the likelihood used (complete, observed, conditional) and the one we worked with here is the *conditional* DIC as categorized by Celeux et *al* (2006). We have avoided using the *observed* DIC because, like the Bayes factor, it is based on evaluating the marginal likelihood whose computation is typically very time-consuming.

4. Simulation study: Finite-sample performance of the QML and BGG estimates

In this Section, a simulation study is undertaken to assess the performance of the QML, BGG Bayes estimates in finite samples.

Concerning finite-properties of the QML and BGG estimates, three instances of the Gaussian $PAR\text{-}SV$ model with period $S = 2$ are considered and are reported respectively in Table 4.1, Table 4.2 and Table 4.3. The parameter $\theta = \left(\alpha_1, \beta_1, \alpha_2, \beta_2, \sigma_1^2, \sigma_2^2\right)'$ are chosen for each instance in order to be in accordance with empirical evidence. In particular, for the three instances the persistence parameter $\beta_1\beta_2$ equals 0.90, 0.95 and 0.99 respectively. We have also set small values for σ_1^2 and σ_2^2 because it is a critical case for the performance of the $QMLE$ as pointed out by Ruiz (1994) and Harvey et *al* (1994) in the standard *SV* case. The choice of $S = 2$ is only motivated by computational and time-consuming considerations. For each instance, we have considered 1000 replications of $PAR\text{-}SV$ series with sample size 1500, for which we calculated the QML and Bayes estimates. Mean of estimates ($\widehat{\theta}_{QML}$ and $\widehat{\theta}_{BGG}$) and their standard deviations (*Std*) over the replications are reported in Tables 4.1-4.3.

For the QML method a non linear optimization routine is required. We have applied a Gauss-Newton type algorithm starting from different values of the θ parameter estimate. For the Bayes Griddy Gibbs estimate, we have taken the same prior distributions for $\omega = (\alpha_1, \beta_1, \alpha_2, \beta_2)'$ across instances:

$$\omega \sim N\left(\omega_0, diag\left(0.05, 0.5, 0.05, 0.5\right)\right), \omega_0 = \left(0, 0, 0, 0\right)',$$
$$\tfrac{1}{\sigma_1^2} \sim \chi_5^2, \quad \tfrac{1}{\sigma_2^2} \sim \chi_5^2,$$

which are quite diffuse, but proper. Concerning initial parameter values, the initial volatility $h^{(0)}$ in the

Gibbs sampler is taken to be the volatility generated by the fitted $GARCH\,(1,1)$, that is $h^{(0)} = h^G$ where

$$\begin{cases} \varepsilon_t = \sqrt{h_t^G}\eta_t \\ h_t^G = \varphi_0 + \varphi_1 \varepsilon_{t-1}^2 + \psi h_{t-1}^G \end{cases}, \quad t \in \mathbb{Z},$$

while the initial log-volatility parameter estimate $\theta^{(0)}$ is taken to be the ordinary least-squares estimate of θ based on the series $\log\left(h^{(0)}\right)$. Furthermore, in the Griddy Gibbs iteration, h_t is generated using 500 grid points and the range of h_t at the l-th Gibbs iteration is taken as in (3.12). Finally, the Gibbs sampler is run for 5500 iterations from which we discarded the first 500 iterations.

	α_1	β_1	α_2	β_2	σ_1	σ_2
True value	-0.5	1	1.2	0.9	0.2	0.3
$QMLE$	-0.5255	1.0728	1.2536	0.9348	0.2629	0.3755
Std	(0.0374)	(0.0691)	(0.0743)	(0.0385)	(0.0926)	(0.0836)
BGG	-0.5004	0.9979	1.1982	0.9061	0.2003	0.2964
Std	(0.0373)	(0.0182)	(0.0421)	(0.0107)	(0.0127)	(0.0165)

Table 4.1: Instance 1- Simulation results for QML and BGG on a Gaussian $PAR\text{-}SV_2$ with $T = 1500$.

	α_1	β_1	α_2	β_2	σ_1	σ_2
True value	-0.5	1	1.2	0.95	0.1	0.2
$QMLE$	-0.5258	1.0799	1.2527	0.9849	0.1394	0.2570
Std	0.0396	0.0587	0.0643	0.0531	0.5697	0.4582
BGG	$-0.4.939$	0.9992	1.2030	0.9505	0.1004	0.2069
Std	0.0133	0.0166	0.0113	0.0105	0.0123	0.0093

Table 4.2: Instance 2- Simulation results for QML and BGG on a Gaussian $PAR\text{-}SV_2$ with $T = 1500$.

	α_1	β_1	α_2	β_2	σ_1	σ_2
True value	-0.5	1	1.2	0.99	0.15	0.1
QMLE	-0.5230	1.3842	1.1773	0.9503	0.0512	0.2608
Std	(0.0733)	(0.0601)	(0.0661)	(0.0642)	(0.4287)	(0.4913)
BGG	-0.5036	1.0025	1.1992	0.9767	0.1606	0.1180
Std	(0.0312)	(0.0150)	(0.0201)	(0.0132)	(0.0101)	(0.0135)

Table 4.3: Instance 3- Simulation results for QML and BGG on a Gaussian $PAR\text{-}SV_2$ with $T = 1500$.

It can be observed that the parameters are quite well estimated by the two methods with an obvious superiority of the Bayes estimate over the $QMLE$. Indeed, in all instances the BGG estimate ($BGGE$) greatly dominates the $QMLE$ in the sense that it has smaller bias and standard deviations. We also observe that the $QMLE$ provides poor estimates as small as the variance parameters σ_1^2 and σ_2^2.

From a theoretical point of view, it would be interesting to compare the $QMLE$ and $BGGE$ when $\log\left(\eta_1^2\right) \sim N(0,1)$, i.e. when $\eta_1 \sim \exp(X/2)$ with $X \sim N(0,1)$. In that case, as emphasized in Section 3, the $QMLE$ reduces to the MLE and it would be more (asymptotically) efficient than the $BGGE$. So through simulations, the $QMLE$ would (in principle) perform better than the $BGGE$ for $PAR\text{-}SV$ series with quite large sample size. However, the BGG method should be adapted to the case of distribution $\eta_1 \sim \exp(X/2)$, which may entails a lot of effort for that distribution ($\exp(N(0,1)/2)$) that seems to have a little interest in practice.

5. Application to the $S\&P$ 500 returns

For the sake of illustration, we propose to fit Gaussian $PAR\text{-}SV$ models (2.1) with various periods to the returns on the $S\&P$ 500 (closing value) index. In order to highlight many possible values of the $PAR\text{-}SV$ period, three types of datasets are considered namely daily, quarterly and monthly $S\&P$ 500 returns. For the three series considered, we use the Bayes Griddy Gibbs estimate thanks to its good finite-sample properties, with number of iterations $M = 5000$ and burn-in 500. As in Section 4, we take the initial volatility $h^{(0)}$ to be the volatility generated by the fitted $GARCH(1,1)$ while the initial log-volatility parameter estimate $\theta^{(0)}$ is taken to be the ordinary least-squares estimate of θ based on the series $\log\left(h^{(0)}\right)$. We have in fact avoided to use the volatility fitted by the periodic $GARCH$ ($PGARCH(1,1)$) model as initial value $h^{(0)}$ because of some numerical difficulties in the corresponding QML estimation when S becomes large (once $S \geq 3$). In the Gibbs step, the volatility $h^{(l)}$ is drawn across $PAR\text{-}SV$ models using the Griddy-Gibbs technique using the same devises as in Section 4, i.e. using 500 grid points and the range of h_t at the l-th Gibbs iteration is taken as in (3.12). All procedures have been applied on a personal computer using Matlab 2013. The *BGG* programs are available from the author upon request.

5.1. Daily $S\&P$ 500 returns: day-of-the-week effect

5.1.1. The data

The first dataset consists of the daily $S\&P$ 500 returns (in decimals) over the sample period starting from January, 01, 2007 to December, 31, 2012, with a total of $T = 1509$ observations. The time series plots of the index (panel (a)) and its return (panel (b)) are presented in Figure 5.1. The same data has also been considered by Chan and Grant (2014).

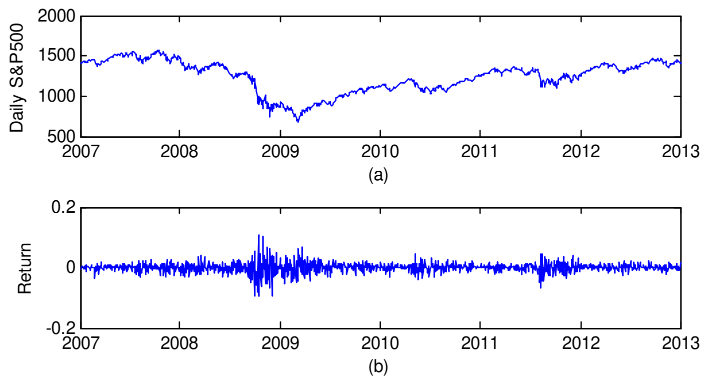

Figure 5.1: Daily $S\&P$ 500 from January 2007 to December 2012.
(a) level, (b) return.

Table 5.1 shows some descriptive statistics for the returns, the absolute returns, the squared return and the log-absolute returns where it may be seen that the data exhibits negative skewness, high kurtosis and low autocorrelation. Moreover, unreported sample correlations with high lags show that the absolute and squared returns are characterized by high persistence with an obvious higher correlation for the absolute returns than the squares. Finally, the log-absolute return looks like a Gaussian much more than do the daily (ε_t), $(|\varepsilon_t|)$ and (ε_t^2). The same finding has been observed by Tsiakas (2006) for the $S\&P$ 500 returns, but for a different sample period.

	Returns (ε_t)	Absolute returns ($\lvert\varepsilon_t\rvert$)	Log-abs. returns ($\log\lvert\varepsilon_t\rvert$)	Squared returns (ε_t^2)
Mean	$4.4711e^{-06}$	0.0102	-5.2239	0.0002
St. Devi.	0.0157	0.0119	1.3095	0.0007
Skewness	-0.2643	2.9540	-0.8449	8.5518
Kurtosis	10.4975	16.5674	4.1138	100.0209
Minimum	-0.0947	0.001	-11.1125	$2.2273e^{-10}$
Maximum	0.1096	0.1096	-2.2112	0.0120
$Corr\,(\varepsilon_t, \varepsilon_{t-1})$	-0.1184	0.2358	0.1250	0.1898
$Corr\,(\varepsilon_t, \varepsilon_{t-2})$	-0.0610	0.3723	0.1915	0.3919
$Corr\,(\varepsilon_t, \varepsilon_{t-3})$	0.0444	0.2896	0.2023	0.1672
$Corr\,(\varepsilon_t, \varepsilon_{t-5})$	-0.0544	0.3865	0.2032	0.3334

Table 5.1: Some descriptive statistics for the daily $S\&P$ 500 returns.

It is by now well documented (Bollerslev and Ghysels, 1996; Franses and Paap, 2000; Tsiakas, 2006) that daily $S\&P$ 500 returns are characterized by the day-of-the-week effect which often suggests the presence of periodicity in volatility with period $S = 5$. While the sample-period chosen here is different from those taken by e.g. Franses and Paap (2000) and Tsiakas (2006) for the same daily $S\&P$ 500 variable, it may be observed from Table 5.2 that the average return and the volatility (approximated by the absolute value) are somewhat different from a day to another. Of course, the difference significancy could be studied more effectively using e.g. the bootstrap approximation of the distribution of the return, along each day as done by Tsiakas (2006). However, this is behind the scoop of this application, which is made only for illustration purposes.

		Sample size	Mean of (ε_t)	Mean of ($\lvert\varepsilon_t\rvert$)	Mean of ($\log\lvert\varepsilon_t\rvert$)	Mean of (ε_t^2)
	Full series	1509	$4.4711e^{-06}$	0.0102	-5.2239	0.0002
1	Monday	284	0.0013	0.0108	-5.1867	0.0003
2	Tuesday	308	-0.0003	0.0099	-5.2355	0.0002
3	Wednesday	311	0.0001	0.0109	-5.0883	0.0003
4	Thursday	305	-0.0002	0.0088	-5.2770	0.0001
5	Friday	301	-0.0007	0.0108	-5.3336	0.0003

Table 5.2: Day of the week effect in the daily $S\&P$ 500 returns.

5.1.2. The models and prior distributions

In order to identify the period of the best fitting PAR-SV model according to the DIC, we estimate six PAR-SV models (2.1) corresponding to each $S \in \{1, ..., 6\}$. For $S = 5$, because of the presence of holidays, model (2.1) in which $\omega_v = \omega_{d(v)}$ and $\sigma_v = \sigma_{d(v)}$ with $d(v) = nS + v$ ($1 \leq v \leq S$, $n \in \mathbb{Z}$) seems not suitable. This is because with model (2.1) each day of a week may have different specification than the same day of the week before. So when $S = 5$ we also estimate the following variant of model (2.1) (henceforth PAR-SV_5^*):

$$\begin{cases} \varepsilon_t = \sqrt{h_t}\eta_t \\ \log(h_t) = \alpha_{d(t)} + \beta_{d(t)} \log(h_{t-1}) + \sigma_{d(t)} e_t \end{cases}, \quad 1 \leq t \leq T, \tag{5.1}$$

in which $d(t)$ is instead defined to be

$$d(t) = \begin{cases} 1 \text{ if the day corresponding to } t \text{ is a Monday} \\ 2 \text{ if the day corresponding to } t \text{ is a Tuesday} \\ \vdots \\ 5 \text{ if the day corresponding to } t \text{ is a Friday.} \end{cases}$$

Such a specification with missing values (see e.g. Franses and Paap, 2000; Regnard and Zakoïan, 2011 in the periodic $GARCH$ case) seems well adapted to explain the day-of-the-week effect.

In calculating the BGG estimate across models, the chosen prior distributions for all the candidate PAR-SV_S models are reported in Table 5.3. These priors are informative, but reasonably flat (cf. Figures 5.2-5.3). When $S = 1$, the prior distributions in Table 5.3 are similar to those proposed by Tsay (2010, Example 12.3) for his SV model. For the variant PAR-SV^* model (5.1), we use the same priors as in the PAR-SV_5 model (2.1). Note that in Table 5.3, the diagonal matrix D_k ($k = 2, 4, 10$) is defined to be

$$D_k(i, j) = \begin{cases} 0 & \text{if} \quad i \neq j \\ 0.05 & \text{if} \quad i = j \text{ is odd} \\ 0.5 & \text{if} \quad i = j \text{ is even} \end{cases}, \quad 1 \leq i, j \leq k. \tag{5.2}$$

$PAR\text{-}SV_S$	Prior for ω : ω_0	$\omega \sim N(\omega_0, \Sigma_0)$ Σ_0	Prior for σ^2 : $a = (a_1, ..., a_S)$	$\frac{a_v \lambda_v}{\sigma_v^2} \sim \chi^2_{a_v}$ $\lambda = (\lambda_1, ..., \lambda_S)$
$S = 1$	$0_{2\times1}$	D_2	5	0.2
$S = 2$	$0_{4\times1}$	D_4	$5 \times \mathbf{1}_2$	$0.2 \times \mathbf{1}_2$
$S = 3$	$0_{6\times1}$	D_6	$5 \times \mathbf{1}_3$	$0.2 \times \mathbf{1}_3$
$S = 4$	$0_{8\times1}$	D_8	$5 \times \mathbf{1}_4$	$0.2 \times \mathbf{1}_4$
$S = 5$	$0_{10\times1}$	D_{10}	$5 \times \mathbf{1}_5$	$0.2 \times \mathbf{1}_5$
$S = 6$	$0_{12\times1}$	D_{12}	$5 \times \mathbf{1}_6$	$0.2 \times \mathbf{1}_6$

Table 5.3: Prior distributions of ω and σ^2 for the candidates $PAR\text{-}SV_S$, $1 \leq S \leq 6$, (D_k, $0_{k\times1}$ and $\mathbf{1}_k$ denote respectively the diagonal matrix given by (5.2), the null vector with k components and the k-vector with all components equal 1).

Using the QML method, the fitted standard $GARCH(1,1)$ specification to the daily $S\&P$ 500 returns is given by

$$h_t^G = \underset{(0.00000081)}{0.00000314} + \underset{(0.0121)}{0.1022}\varepsilon_{t-1}^2 + \underset{(0.0138)}{0.8822}h_{t-1}^G,$$

with standard deviations of estimates in parentheses. The found volatility h^G is used to initialize the volatility parameter $h^{(0)}$ in the Gibbs sampler across all estimated $PAR\text{-}SV_S$ (and $PAR\text{-}SV_5^*$) models.

5.1.3. Results

The estimated DIC's across $PAR\text{-}SV$ models, their computation times (in minutes), their numerical standard errors (approximated by their standard deviations over G replications) and the monodromy parameters of all estimated models are reported in Table 5.4. In computing the standard errors of DIC, we have replicated the BGG procedure (Algorithm 3.1) $G = 500$ times.

	$PAR\text{-}SV_1$	$PAR\text{-}SV_2$	$PAR\text{-}SV_3$	$PAR\text{-}SV_4$	$PAR\text{-}SV_5$	$PAR\text{-}SV_5^*$	$PAR\text{-}SV_6$
$-DIC$ (Std)	8872.1155 (1.0737)	8869.2121 (1.1277)	8863.8238 (1.2547)	8866.6620 (0.9992)	8864.5966 (1.1825)	**8881.8162** (**0.8261**)	8868.5853 (1.2823)
Rank	2	3	7	5	6	**1**	4
Time	80.3535	81.9259	81.5512	82.7189	82.9024	**82.7828**	83.5396
Monod.	0.8509	0.8332	0.7277	0.6423	0.5918	**0.5853**	0.5160

Table 5.4: Estimated DIC, standard deviation, ranking, computation time (in minutes) and the monodromy (Monod.) estimate for the candidate $PAR\text{-}SV_S$ ($1 \leq S \leq 6$) and $PAR\text{-}SV_5^*$ models.

From Table 5.4, some broad conclusions are in order. Firstly, the DIC's corresponding to the $PAR\text{-}SV_S$ $(1 \leq S \leq 6)$ models given by (2.1) are very close to each other. So, with regard to the standard errors of the DIC's, which are reasonably small, it is difficult to distinguish between the corresponding $PAR\text{-}SV_S$ $(1 \leq S \leq 6)$ models despite the inherent ranking reported in Table 5.4. On the other hand, the DIC favors the $PAR\text{-}SV_5^*$ given by (5.1), whose value (-8881.8162) is quite small than the others. Secondly, while the BGG method is relatively time-consuming for all $PAR\text{-}SV$ models, a fact that is well known in the single move approach, the computation time is almost similar across $PAR\text{-}SV$ models in spite of the increasing number of parameters when S tends to be large. So fitting a periodic $PAR\text{-}SV$ model is carried out without increasing computational cost compared to the non-periodic SV. Thirdly, the monodromy parameters $\prod_{v=1}^{S} \beta_v$ across models are quite large, which suggests a strong persistence in volatility.

According to the DIC, the best model is the $PAR\text{-}SV_5^*$ given by (5.1), whose parameters, their $MCMC$ standard deviations (std), their NSE and their RNI are reported in Table 5.6. As a benchmark, Table 5.5 reports the same information concerning the second best model ranked by the DIC, which is the standard SV corresponding to $PAR\text{-}SV_1$. Due to lack of space the remaining estimated models are not presented here, but are available from the author. Further, prior and posterior distributions of the estimates for the $PAR\text{-}SV_1$ and $PAR\text{-}SV_5^*$ are plotted in Figure 5.2 and Figure 5.3 respectively.

$PAR\text{-}SV_1$ *parameter*	Posterior mean	$MCMC$ *std*	NSE	RNI
α	-1.2822	0.1132	0.0041	0.1431
β	0.8509	0.0133	0.0004	0.1389
σ^2	0.2535	0.0128	0.0006	0.2525

Table 5.5: BGG parameter estimates for the $PAR\text{-}SV_1$ (standard SV).

Day	$PAR\text{-}SV_5^*$ parameter	Posterior mean	$MCMC$ std	NSE	RNI
Monday	α_1	-0.5655	0.1014	0.0026	0.7179
	β_1	0.9351	0.0121	0.0003	0.4708
	σ_1^2	0.2482	0.0270	0.0004	0.2020
Tuesday	α_2	-1.0438	0.0942	0.0022	0.6025
	β_2	0.8795	0.0116	0.0003	0.5675
	σ_2^2	0.2471	0.0268	0.0003	0.1126
Wednesday	α_3	-0.8723	0.0986	0.0021	0.5202
	β_3	0.9010	0.0119	0.0003	0.7461
	σ_3^2	0.2348	0.0241	0.0003	0.1457
Thursday	α_4	-1.0149	0.0969	0.0016	0.2827
	β_4	0.8835	0.0116	0.0002	0.3509
	σ_4^2	0.2416	0.0253	0.0009	1.3812
Friday	α_5	-0.9366	0.0992	0.0017	0.3159
	β_5	0.8938	0.0119	0.0002	0.2315
	σ_5^2	0.2336	0.0252	0.0005	0.4316

Table 5.6: BGG parameter estimates for the $PAR\text{-}SV_5^*$.

It may be seen from Table 5.5-5.6 that the parameters appear quite well estimated as shown by their low $MCMC$ standard deviations, low RNI and small NSE. The latter clearly shows that even with the single move approach, when a suitable choice of the range of h in the Griddy-Gibbs procedure is made, the $MCMC$ estimates mixe well. This is confirmed by the low autocorrelations of the estimates (cf. Figure 5.4). Moreover, from Table 5.6 it can be observed that the parameters are quite different from a day to another, especially for the α_v and β_v $(1 \leq v \leq 5)$. On the other hand, the estimates are comparable with similar models in the literature when $S = 1$. Prior and posterior distributions of the estimates for the $PAR\text{-}SV_1$ and $PAR\text{-}SV_5^*$ are plotted in Figure 5.2 and Figure 5.3 respectively. The prior distributions used are, as pointed out above, relatively noninformative while the posterior distributions are quite concentrated. In addition, from Figure 5.5 the volatilities induced by the $GARCH(1,1)$ (dashed-line) and $PAR\text{-}VS_5^*$ (solid line) models have similar pattern. Note finally that these result were quite stable to using different initial values, priors, and numbers of iterations for the Gibbs sampler. However, the efficiency of the Gibbs sampler greatly depends on the choice of the range of h in the Griddy-Gibbs step.

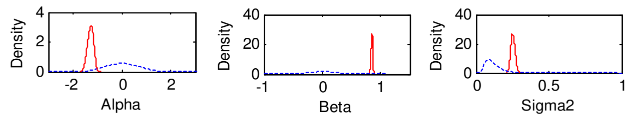

Figure 5.2: Prior (dashed line) and *BGG* posterior (solid line) distributions of parameters in the $PAR\text{-}SV_1$ (aperiodic *SV*) model.

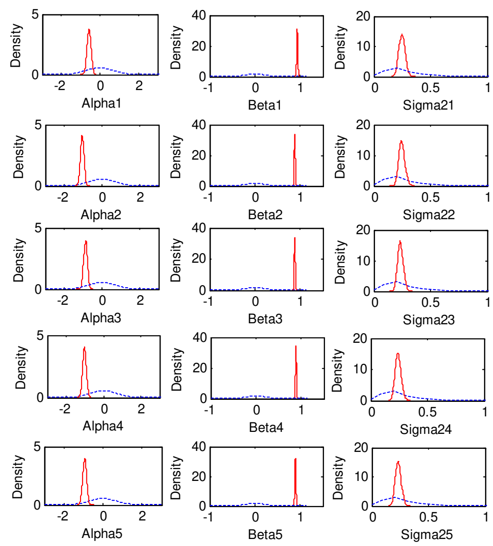

Figure 5.3: Prior (dashed line) and *BGG* posterior (solid line) distributions of parameters in the $PAR\text{-}SV_5^*$ model.

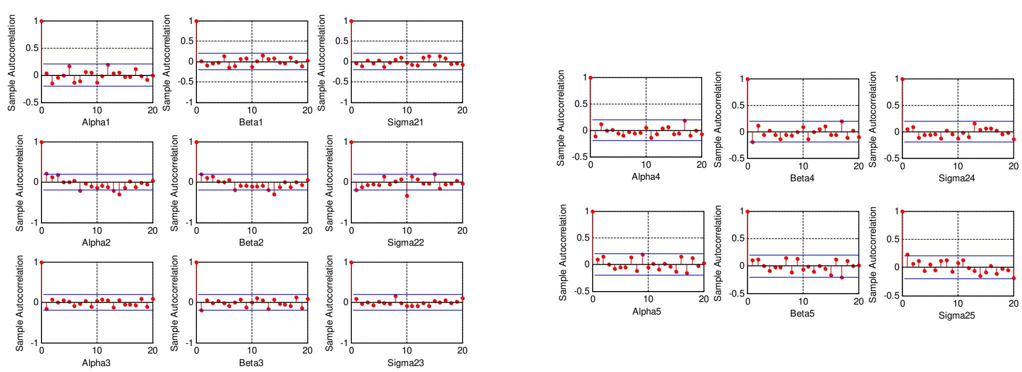

Figure 5.4: Sample autocorrelations of the $PAR\text{-}SV_5^*$ parameter estimates.

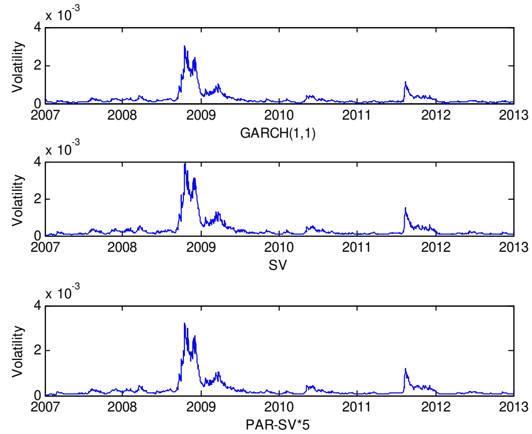

Figure 5.5: Volatilities induced by the $GARCH\,(1,1)$,

the SV and $PAR\text{-}SV_5^*$.

5.2. Quarterly $S\&P$ 500 returns

The second dataset consists of the quarterly $S\&P$ 500 returns over the sample period from the first quarter ($Q1$) 1871 to the fourth quarter ($Q4$) 2012, with a total of $T = 567$ observations. The index is calculated by taking average price per share in month ending quarter. The time series plots of the index series and its return are displayed in Figure 5.6. The data are given from Shiller (2015).

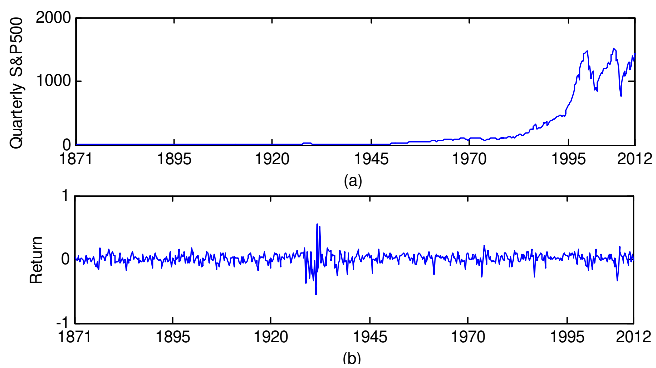

Figure 5.6: Quarterly $S\&P$ 500 index : (a) level and (b) return.

We estimated five $PAR\text{-}SV$ models (2.1) corresponding to each $S \in \{1, ..., 5\}$ using the same prior distributions as in Table 5.3 (for $1 \leq v \leq 5$). The estimated volatility via the $GARCH(1,1)$ model, which is used to initialize the volatility in the Gibbs sampler, has the following specification

$$h_t^G = \underset{(0.0035)}{0.0010} + \underset{(0.0577)}{0.1792}\varepsilon_{t-1}^2 + \underset{(0.0674)}{0.6796} h_{t-1}^G,$$

with standard errors of estimates in parentheses. The estimated DIC's across $PAR\text{-}SV$ models, their computation times (in minutes), their numerical standard errors and the corresponding monodromy parameters are reported in Table 5.7. The standard errors of DIC are calculated as above.

	$PAR\text{-}SV_1$	$PAR\text{-}SV_2$	$PAR\text{-}SV_3$	$PAR\text{-}SV_4$	$PAR\text{-}SV_5$
DIC (Std)	-1206.4619 (2.0577)	-1210.4912 (2.2225)	-1209.6398 (1.9911)	$\mathbf{-1211.9735}$ (**1.0370**)	-1207.6028 (2.2684)
Rank	5	2	3	**1**	4
Time	29.7357	30.2224	31.3725	**30.3580**	32.4587
Monod.	0.6298	0.5102	0.4476	**0.4304**	0.3841

Table 5.7: Estimated DIC, standard deviations, ranking, computation time (in minutes) and the monodromy (Monod.) estimates for the candidate $PAR\text{-}SV_S$ $(1 \leq S \leq 5)$ models.

From Table 5.7, the DIC selects the four-periodic $PAR\text{-}SV_4$ with smallest value $\mathbf{-1211.9735}$. Such a value is not so far from those of the remaining $PAR\text{-}SV_S$ models ($S \neq 4$) regarding their numerical standard errors. On the other hand, the corresponding computation times are quite comparable while the monodromy parameters are less important than in the daily return case. The parameters of the found $PAR\text{-}SV_4$ model, their $MCMC$ standard deviations, their NSE and their RNI are listed in Table 5.8.

Quarter	$PAR\text{-}SV_4$ $parameters$	Posterior mean	$MCMC$ std	NSE	RNI
	α_1	-1.0552	0.0693	0.0007	0.5029
$Q1$	β_1	0.7861	0.0183	0.0002	0.7718
	σ_1^2	0.2283	0.0359	0.0005	1.1526
	α_2	-0.7527	0.0692	0.0006	0.4521
$Q2$	β_2	0.8379	0.0187	0.0003	1.2665
	σ_2^2	0.2752	0.0436	0.0009	2.3623
	α_3	-0.8548	0.0709	0.0012	1.5954
$Q3$	β_3	0.8181	0.0189	0.0003	1.4664
	σ_3^2	0.2547	0.0397	0.0007	1.5453
	α_4	-0.9423	0.0697	0.0011	1.3967
$Q4$	β_4	0.7987	0.0188	0.0003	1.8023
	σ_4^2	0.2766	0.0438	0.0009	2.1791

Table 5.8: BGG parameter estimates for the selected $PAR\text{-}SV_4$ model.

The same conclusions for the daily return case may be obtained: the parameters are quite well estimated in view of their low standard deviations, low RNI and small NSE (cf. Table 5.8). Moreover, the posterior distributions are fairly concentrated (cf. Figure 5.7). On the other hand, the parameters are quite different from a quarter to another especially for the α_v and β_v. However, in overall, the estimates seem slightly less accurate than in the daily return case, which is perhaps due to the smaller sample size. Finally, Figure 5.8 plots the volatilities generated by the $GARCH(1,1)$ (panel (a)) and the $PAR\text{-}SV_4$ (panel (b)) where it may be seen that they display a very similar pattern.

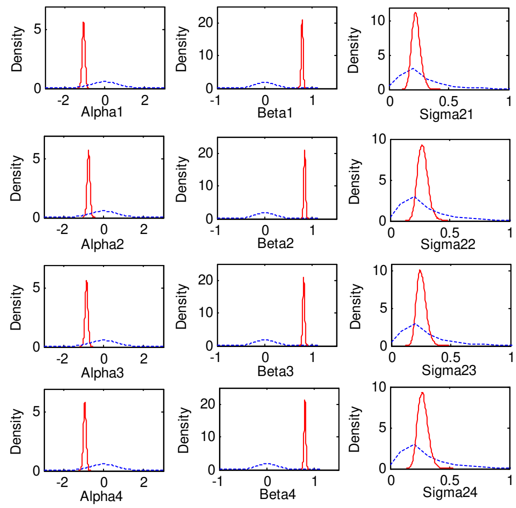

Figure 5.7: Prior (dashed line) and posterior (solid line) distributions of parameters in the $PAR\text{-}SV_4$ model.

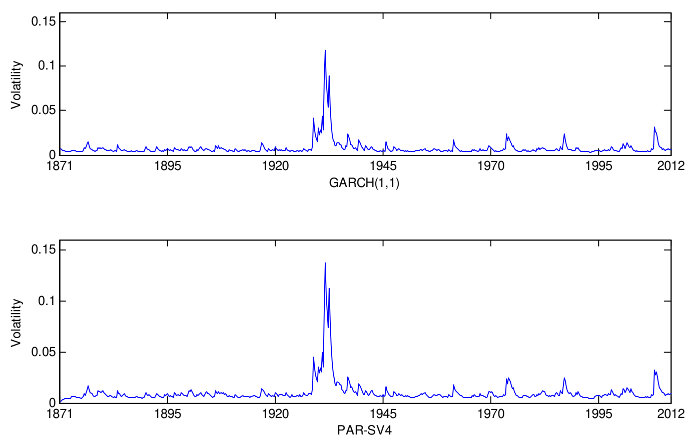

Figure 5.8: Volatilities induced by the $GARCH\,(1,1)$ and $PAR\text{-}SV_4$.

5.3. Monthly $S\&P$ 500 returns

The third dataset consists of the return of the monthly $S\&P$ 500 index from January 1950 to January 2015, involving 780 observations. The returns are computed using the first adjusted closing index of each month. Plots of the $S\&P$ 500 index and its return are given in Figure 5.9. A similar monthly series, but on different sample period has been studied by Tsay (2010, example 12.3) via a SV model.

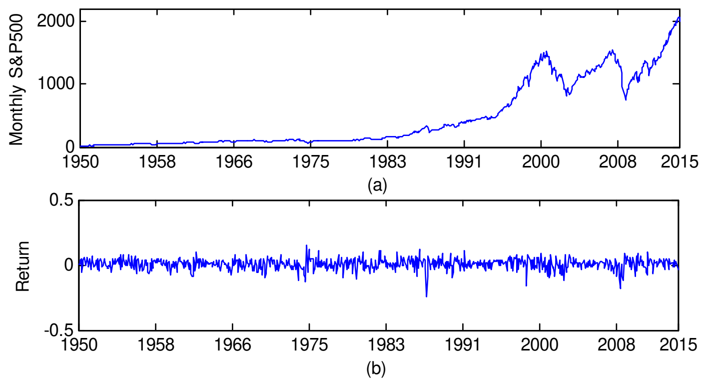

Figure 5.9: Monthly $S\&P$ 500 index: (a) level and (b) return.

We estimated twelve PAR-SV models (2.1) corresponding to each $S \in \{1, ..., 12\}$ using the prior distributions presented in Table 5.9.

$PAR\text{-}SV_S$	Prior for ω : ω_0	$\omega \sim N(\omega_0, \Sigma_0)$ Σ_0	Prior for σ^2 : $a = (a_1, ..., a_S)$	$\frac{a_v \lambda_v}{\sigma_v^2} \sim \chi^2_{a_v}$ $\lambda = (\lambda_1, ..., \lambda_S)$
$S = 1$	$0_{2\times 1}$	D_2	5	0.2
$S = 2$	$0_{4\times 1}$	D_4	$5 \times \mathbf{1}_2$	$0.2 \times \mathbf{1}_2$
$S = 3$	$0_{6\times 1}$	D_6	$5 \times \mathbf{1}_3$	$0.2 \times \mathbf{1}_3$
$S = 4$	$0_{8\times 1}$	D_8	$5 \times \mathbf{1}_4$	$0.2 \times \mathbf{1}_4$
$S = 5$	$0_{10\times 1}$	D_{10}	$5 \times \mathbf{1}_5$	$0.2 \times \mathbf{1}_5$
$S = 6$	$0_{12\times 1}$	D_{12}	$5 \times \mathbf{1}_6$	$0.2 \times \mathbf{1}_6$
$S = 7$	$0_{14\times 1}$	D_{14}	$5 \times \mathbf{1}_7$	$0.2 \times \mathbf{1}_7$
$S = 8$	$0_{16\times 1}$	D_{16}	$5 \times \mathbf{1}_8$	$0.2 \times \mathbf{1}_8$
$S = 9$	$0_{18\times 1}$	D_{18}	$5 \times \mathbf{1}_9$	$0.2 \times \mathbf{1}_9$
$S = 10$	$0_{20\times 1}$	D_{20}	$5 \times \mathbf{1}_{10}$	$0.2 \times \mathbf{1}_{10}$
$S = 11$	$0_{22\times 1}$	D_{22}	$5 \times \mathbf{1}_{11}$	$0.2 \times \mathbf{1}_{11}$
$S = 12$	$0_{24\times 1}$	D_{24}	$5 \times \mathbf{1}_{12}$	$0.2 \times \mathbf{1}_{12}$

Table 5.9: Prior distributions of ω and σ^2 for the candidates $PAR\text{-}SV_S$, $1 \le S \le 12$, (D_k, $0_{k\times 1}$ and $\mathbf{1}_k$ denote respectively the diagonal matrix given by (5.2), the null vector with k components and the k-vector with all components equal 1).

The volatility generated by the $GARCH(1,1)$ model, which is used to initialize the volatility in the Gibbs sampler across estimated $PAR\text{-}SV_S$ models, is given by

$$h_t^G = \underset{(0.00002)}{0.0001} + \underset{(0.0245)}{0.1058}\varepsilon_{t-1}^2 + \underset{(0.0271)}{0.8502}h_{t-1}^G,$$

with standard errors of estimates in parentheses. The estimated DIC's for $PAR\text{-}SV$ models, their computation times (in minutes), their numerical standard errors and the corresponding monodromy parameters are reported in Table 5.7. The standard errors of DIC are calculated using 500 replications of the BGG

procedure.

	$PAR\text{-}SV_1$	$PAR\text{-}SV_2$	$PAR\text{-}SV_3$	$PAR\text{-}SV_4$	$PAR\text{-}SV_5$	$PAR\text{-}SV_6$
DIC (Std)	-2685.1845 (1.2577)	-2685.5848 (1.8528)	-2682.9659 (2.2164)	-2679.9444 (2.0055)	-2684.1986 (1.8917)	-2684.4897 (0.9922)
Rank	6	4	11	12	9	7
Time	41.9776	42.6814	42.5118	43.0081	42.7517	44.1573
Monod.	0.6259	0.7249	0.7593	0.6124	0.6255	0.5969

	$PAR\text{-}SV_7$	$PAR\text{-}SV_8$	$PAR\text{-}SV_9$	$PAR\text{-}SV_{10}$	$PAR\text{-}SV_{11}$	$PAR\text{-}SV_{12}$
DIC (Std)	-2685.6729 (2.0248)	-2684.4211 (1.9524)	-2685.5062 (2.0473)	-2684.03953 (2.1246)	-2685.9018 (1.6524)	$\mathbf{-2686.6698}$ **(1.2684)**
Rank	3	8	5	10	2	**1**
Time	43.3544	43.9715	43.2856	43.5522	44.3780	**44.9834**
Monod.	0.5726	0.5399	0.5335	0.5099	0.4973	**0.4771**

Table 5.10: Estimated DIC, standard deviations, ranking, computation time (in minutes) and the monodromy (Monod.) estimates for the candidate $PAR\text{-}SV_S$ $(1 \leq S \leq 12)$ models.

According to the DIC, the best model is the 12-periodic $PAR\text{-}SV_{12}$ with value -2686.6698. However, the DIC's in Table 5.10 are very close to each other, so in view of their standard errors, it is difficult to discriminate between the corresponding models. On the other hand, as in the quarterly return case, the monodromy estimates are around half a unity while the computation times are close to each other. The specification of the selected $PAR\text{-}SV_{12}$ model is given in Table 5.11.

	θ	Post. mean	MC std	NSE	RNI		θ	Post. mean	MC std	NSE	RNI
	α_1	-0.6197	0.1264	0.0058	0.0727		α_7	-0.3703	0.0738	0.0103	0.2141
Jan.	β_1	0.8845	0.0367	0.0007	0.0399	Jul.	β_7	0.9364	0.0418	0.0020	0.2514
	σ_1^2	0.2287	0.0442	0.0033	0.6241		σ_7^2	0.2485	0.0644	0.0049	0.6491
	α_2	-1.2150	0.1071	0.0123	0.3904		α_8	-0.3928	0.0980	0.0191	0.8508
Feb.	β_2	0.7989	0.0330	0.0019	0.3815	Aug.	β_8	0.9237	0.0386	0.0030	0.6539
	σ_2^2	0.2364	0.0531	0.0016	0.1063		σ_8^2	0.2829	0.0631	0.0032	0.2822
	α_3	0.4881	0.1057	0.0079	0.1253		α_9	-0.4057	0.0860	0.0196	0.8343
Mar.	β_3	1.0821	0.0413	0.0011	0.0799	Sep.	β_9	0.9250	0.0379	0.0039	1.1715
	σ_3^2	0.3095	0.0796	0.0065	0.7369		σ_9^2	0.2443	0.0508	0.0021	0.1871
	α_4	0.3656	0.0719	0.0093	0.2148		α_{10}	-0.3348	0.1953	0.0135	0.5348
Apr.	β_4	1.0500	0.0373	0.0022	0.3884	Oct.	β_{10}	0.9400	0.0347	0.0034	1.0388
	σ_4^2	0.2841	0.0778	0.0064	0.7499		σ_{10}^2	0.3017	0.0736	0.0063	0.8232
	α_5	-0.2547	0.0863	0.0173	0.6454		α_{11}	-0.4065	0.0937	0.0081	0.1240
May	β_5	0.9572	0.0381	0.0034	0.8705	Nov.	β_{11}	0.9259	0.0422	0.0024	0.3609
	σ_5^2	0.2552	0.0596	0.0077	1.8386		σ_{11}^2	0.2768	0.0676	0.0054	0.7129
	α_6	-0.3875	0.0977	0.0305	1.8303		α_{12}	-0.4025	0.0835	0.0182	0.7401
Jun.	β_6	0.9352	0.0408	0.0054	1.9583	Dec.	β_{12}	0.9351	0.0403	0.0013	0.1194
	σ_6^2	0.2665	0.0504	0.0019	0.1541		σ_{12}^2	0.2581	0.0520	0.0046	0.8719

Table 5.11: BGG parameter estimates for the selected $PAR\text{-}SV_{12}$ model.

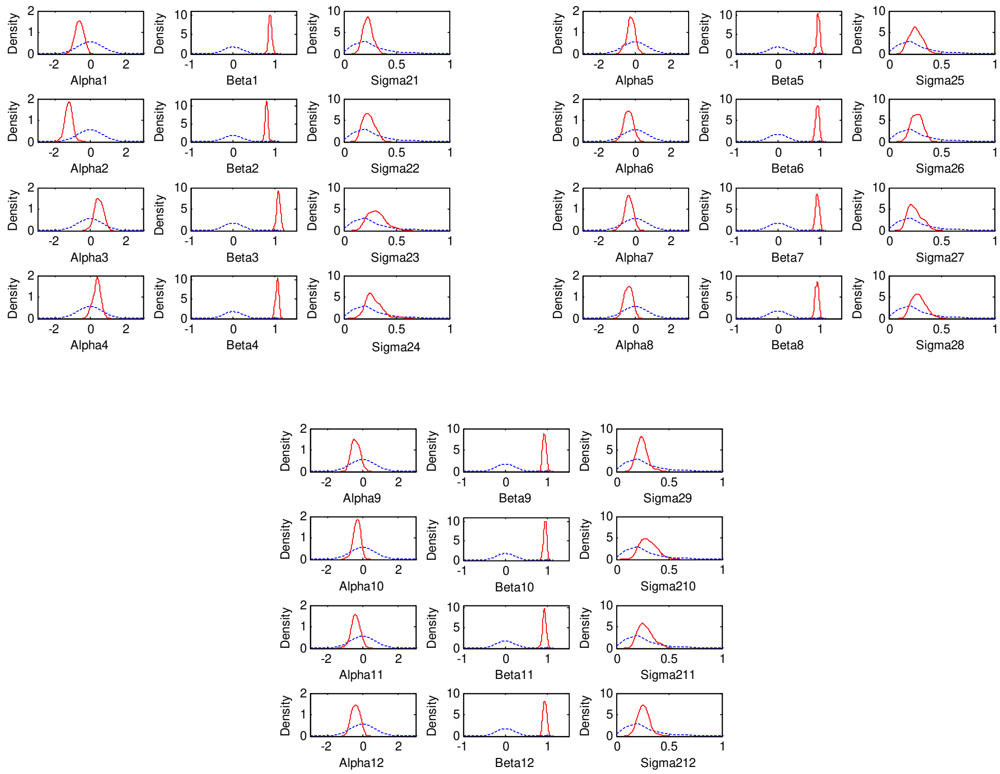

Figure 5.10: Prior (dashed line) and posterior (solid line) distributions of parameters in the $PAR\text{-}SV_{12}$ model.

From Table 5.11 and Figure 5.10, it may be concluded that the estimates are relatively good in spite of the small sample size compared to the large number of parameters to estimate. The $MCMC$ std, the RNI and the NSE are fairly low while the posterior distributions are quite concentrated. Moreover, the parameters seem different from a month to another, especially for the α_v and β_v. Finally, from Figure 5.11, the volatilities induced by the $GARCH(1,1)$ (upper panel) and the 12-periodic $PAR\text{-}SV_{12}$ (lower panel) have a similar behavior in both shape and magnitude.

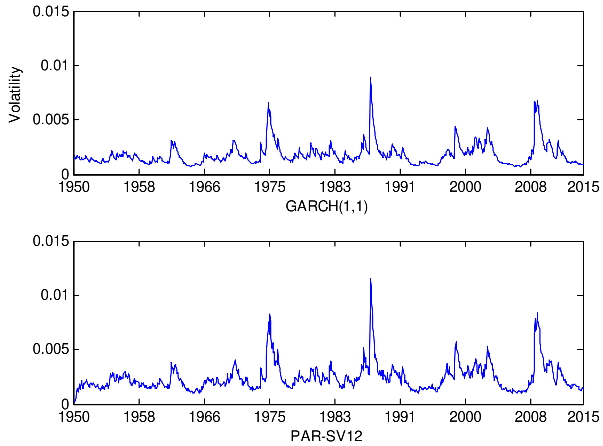

Figure 5.11: Volatilities induced by the $GARCH\,(1,1)$ and $PAR\text{-}SV_{12}$.

6. Conclusion

In this paper we have proposed a stochastic volatility model whose log-volatility follows a periodic autoregression. This model may be seen as a flexible complementary to the periodic $GARCH$ process because it overcomes the limitation that the volatility is only driven by the past of the process. Moreover, the periodic time-variation of the parameters allows a more flexible periodic volatility modelling compared to the time-invariant seasonal SV or deterministic periodic SV. As we have seen, statistical inference for this model may be easily done using the Bayesian $MCMC$ approach without additional computational cost compared to the standard SV case. While the $PAR\text{-}SV$ model allows modelling some financial features such as periodicity in volatility, volatility clustering and excess kurtosis, it seems incapable of representing other observed facts. In particular, excess kurtosis implied by the model might be only of a given order of magnitude and may be less than kurtosis generated by heavy tail innovation η_1, like the Student distribution. So, various interesting generalizations of the proposed $PAR\text{-}SV$ model to account for additional features like large excess kurtosis, leverage effect, change in volatility regime and simultaneous volatility dependence are needed and may constitute future research. In particular, $PAR\text{-}SV$ with heavy tailed innovations like a student or mixture Gaussian distribution, Markov switching $PAR\text{-}SV$, $PAR\text{-}SV$ models with correlated error terms, and multivariate versions of the $PAR\text{-}SV$ are appealing. Finally, a multi-move $MCMC$ approach for estimating $PAR\text{-}SV$ models would be of great interest.

Acknowledgements I am deeply grateful to an Associate Editor and two Referees for their careful reading, helpful comments and useful suggestions that have greatly helped me in improving substantially the earlier version of the manuscript.

References

[1] Aknouche, A. and Al-Eid, E. (2012). Asymptotic inference of unstable periodic *ARCH* processes. *Statistical Inference for Stochastic Processes*, **15**, 61-79.

[2] Aknouche, A. and Bibi, A. (2009). Quasi-maximum likelihood estimation of periodic *GARCH* and periodic *ARMA-GARCH* processes. *Journal of Time series Analysis*, **30**, 19-46.

[3] Berg, A., Meyer, R. and Yu, J. (2004). Deviance information criterion for comparing stochastic volatility models. *Journal of Business and Economic Statistics*, **22**, 107-120.

[4] Bollerslev, T. and Ghysels, E. (1996). Periodic autoregressive conditional heteroskedasticity. *Journal of Business and Economic Statistics*, **14**, 139-151.

[5] Box, G.E.P. and Tiao, G.C. (1973). *Bayesian Inference in Statistical Analysis*. Addison-Wesley: Reading, MA.

[6] Breidt, F.J. (1997). A threshold autoregressive stochastic volatility. Preprint, http://citeseer.ist.psu.edu/103850.html.

[7] Breidt, F.J., Crato, N. and De Lima, P. (1998). The detection and estimation of long memory in stochastic volatility. *Journal of Econometrics*, **83**, 325-348.

[8] Brockwell, P.J. and Davis, R.A. (1991). *Time series: Theory and methods*, 2nd edt. New York: Springer.

[9] Carvalho, C.M. and Lopes, H.F. (2007). Simulation-based sequential analysis of Markov switching stochastic volatility models. *Computational Statistics & Data Analysis*, **51**, 4526-4542.

[10] Celeux, G., Forbes, F., Robert, C.P. and Titterington, D.M. (2006). Deviance information criteria for missing data models. *Bayesian Analysis*, **1**, 651-674.

[11] Chan, C.C. and Grant, A.L. (2014). Issues in comparing stochastic volatility models using the deviance information criterion. *CAMA Working Paper 51/2014. Australian National University.*

[12] Chib, S., Nardari, F. and Shephard, N. (2002). Markov chain Monte Carlo methods for stochastic volatility models. *Journal of Econometrics*, **108**, 281-316.

[13] Davis, R.A. and Mikosch, T. (2009). Probabilistic properties of stochastic volatility models. In "*Handbook of financial time series*", Thomas Mikosch, (Ed.). Part 2, 255-267, Springer.

[14] Dunsmuir, W. (1979). A Central Limit Theorem for parameter estimation in stationary vector time series and its applications to models for a signal observed with noise. *Annals of Statistics*, **7**, 490-506.

[15] Francq, C. and Zakoïan, J.M. (2010). *GARCH Models: Structure, Statistical Inference and Financial Applications*. Wiley, New York.

[16] Francq, C. and Zakoïan, J.M. (2006). Linear-representation bases estimation of stochastic volatility models. *Scandinavian Journal of Statistics*, **33**, 785-806.

[17] Geweke, J. (1989). Bayesian inference in econometric models using Monte Carlo integration. *Econometrica*, **57**, 1317-1339.

[18] Geyer, C.J. (1992). Practical Markov Chain Monte Carlo. *Statistical Science*, **7**, 473-511.

[19] Ghysels, E., Harvey, A.C. and Renault, E. (1996). Stochastic volatility. In: "*Statistical Methods in Finance*", Rao, C.R., Maddala, G.S. (ed.). North-Holland, Amsterdam.

[20] Ghysels, E. and Osborn, D. (2001). *The econometric analysis of seasonal time series*. Cambridge University Press.

[21] Gladyshev, E.G. (1961). Periodically correlated random sequences. *Soviet Mathematics*, **2**, 385-388.

[22] Harvey, A.C., Ruiz, E. and Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, **61**, 247-264.

[23] Jacquier, E., Polson, N.G. and Rossi, P.E. (1994). Bayesian analysis of stochastic volatility models. *Journal of Business and Economic Statistics*, **12**, 413-417.

[24] Kim, S., Shephard, N. and Chib, S. (1998). Stochastic volatility: likelihood inference and comparison with *ARCH* models. *Review of Economic Studies*, **65**, 361-93.

[25] Koopman, S.J., Ooms, M. and Carnero, M.A. (2007). Periodic seasonal *Reg-ARFIMA-GARCH* models for daily electricity spot prices. *Journal of the American Statistical Association*, **102**, 16-27.

[26] Ljung, L. and Söderström, T. (1983). *Theory and practice of recursive identification*. MIT Press, Cambridge, Massachusetts.

[27] Meyn, S. and Tweedie, R. (2009). *Markov chains and stochastic stability*. Springer Verlag, New York, 2nd edition.

[28] Nakajima, J. and Omori, Y. (2009). Leverage, heavy-tails and correlated jumps in stochastic volatility models. *Computational Statistics & Data Analysis*, **53**, 2335-2353.

[29] Omori, Y., Chib, S., Shephard, N. and Nakajima, J. (2007). Stochastic volatility with leverage: fast and efficient likelihood inference. *Journal of Econometrics*, **140**, 425-49.

[30] Osborn, D.R., Savva, C.S. and Gill, L. (2008). Periodic dynamic conditional correlations between stock markets in Europe and the US. *Journal of Financial Econometrics*, **6**, 307-325.

[31] Regnard, N. and Zakoïan, J.M. (2011). A conditionally heteroskedastic model with time-varying coefficients for daily gas spot prices. *Energy Economics*, **33**, 1240-1251.

[32] Ritter, C. and Tanner, M.A. (1992). Facilitating the Gibbs sampler: the Gibbs stopper and the Griddy-Gibbs sampler. *Journal of the American Statistical Association*, **87**, 861-870.

[33] Ruiz, E. (1994). Quasi-maximum likelihood estimation of stochastic variance models. *Journal of Econometrics*, **63**, 284-306.

[34] Sakai, H. and Ohno, S. (1997). On backward periodic autoregressive processes. *Journal of Time Series Analysis*, **18**, 415-427.

[35] Shiller, R.J. (2015). *Irrational exuberance*. 3rd edition, Princeton University Press.

[36] Sigauke, C. and Chikobvu, D. (2011). Prediction of daily peak electricity demand in South Africa using volatility forecasting models. *Energy Economics*, **33**, 882-888.

[37] So, M.K.P., Lam, K., and Li, W.K. (1998). A stochastic volatility model with Markov switching. *Journal of Business and Economic Statistics*, **16**, 244-253.

[38] Spiegelhalter, D.J., Best, N.G., Carlin, B.P. and van der Linde, A. (2002). Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society*, **B64**, 583-639.

[39] Taylor, J.W. (2006). Density forecasting for the efficient balancing of the generation and consumption of electricity. *International Journal of Forecasting*, **22**, 707-724.

[40] Taylor, S.J. (1982). Financial returns modelled by the product of two stochastic processes- a study of daily sugar prices, 1961-1979. *In "Time Series Analysis: Theory and Practice"*, Anderson O.D. (ed.). North-Holland, Amsterdam, 203-226.

[41] Tiao, G.C. and Grupe, M.R. (1980). Hidden periodic autoregressive-moving average models in time series data. *Biometrika*, **67**, 365-373.

[42] Tsay, R.S. (2010). *Analysis of financial time series: Financial Econometrics*. 3rd edition, Wiley.

[43] Tsiakas, I. (2006). Periodic stochastic volatility and fat tails. *Journal of Financial Econometrics*, **4**, 90-135.

[44] Zhu, L. and Carlin, B.P. (2000). Comparing hierarchical models for spatio-temporally misaligned data using the Deviance Information Criterion. *Statistics in Medicine*, **19**, 2265-2278.