

Um Estudo sobre Modelos para Volatilidade Estocástica

Dissertação de Mestrado

André Silva de Queiroz

Orientadora: Prof.^a Dr.^a Cibele Queiroz da Silva

Departamento de Estatística
Universidade de Brasília

10 de dezembro de 2015



Introdução

- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.
- Modelos Dinâmicos.



Introdução

- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.
- Modelos Dinâmicos.



Introdução

- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.
- Modelos Dinâmicos.



Introdução

- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.
- Modelos Dinâmicos.



Introdução

- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.
- Modelos Dinâmicos.



Introdução

- A importância dos modelos de séries temporais.
- Dados com a variância variável:
 - Tradicionalmente: ARCH e GARCH.
 - Nova proposta: Modelos de Volatilidade Estocástica.
- Inferência Bayesiana.
- Modelos Dinâmicos.



Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:

- Equação das Observações:

$$Y_t = F_t' \theta_t + \nu_t, \quad \nu_t \sim \mathcal{N}(0, V_t), \quad (1)$$

- Equação do Sistema:

$$\theta_t = G_t \theta_{t-1} + \omega_t, \quad \omega_t \sim \mathcal{N}(0, W_t). \quad (2)$$

- O MLD é caracterizado pela quádrupla $\{F_t, G_t, V_t, W_t\}$.



Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:

- Equação das Observações:

$$Y_t = \mathbf{F}_t' \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim \mathcal{N}(0, V_t), \quad (1)$$

- Equação do Sistema:

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t). \quad (2)$$

- O MLD é caracterizado pela quádrupla $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$.



Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:
 - **Equação das Observações:**

$$Y_t = \mathbf{F}_t' \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim \mathcal{N}(0, V_t), \quad (1)$$

- **Equação do Sistema:**

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \omega_t, \quad \omega_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t). \quad (2)$$

- O MLD é caracterizado pela quádrupla $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$.



Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:

- **Equação das Observações:**

$$Y_t = \mathbf{F}_t' \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim \mathcal{N}(0, V_t), \quad (1)$$

- **Equação do Sistema:**

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t). \quad (2)$$

- O MLD é caracterizado pela quádrupla $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$.



Modelo Linear Dinâmico (MLD)

- Família mais simples dentre os modelos dinâmicos.
- Segundo West e Harrison (1997) a definição do modelo linear dinâmico univariado é:

- **Equação das Observações:**

$$Y_t = \mathbf{F}_t' \boldsymbol{\theta}_t + \nu_t, \quad \nu_t \sim \mathcal{N}(0, V_t), \quad (1)$$

- **Equação do Sistema:**

$$\boldsymbol{\theta}_t = \mathbf{G}_t \boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \quad \boldsymbol{\omega}_t \sim \mathcal{N}(\mathbf{0}, \mathbf{W}_t). \quad (2)$$

- O MLD é caracterizado pela quádrupla $\{\mathbf{F}_t, \mathbf{G}_t, V_t, \mathbf{W}_t\}$.



- Se $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}$, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:
 - Informação Inicial:

$$(\theta_0 | D_0) \sim \mathcal{N}(m_0, C_0). \quad (3)$$



- Se $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}$, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:
 - Informação Inicial:

$$(\theta_0|D_0) \sim \mathcal{N}(m_0, C_0). \quad (3)$$



- Se $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}$, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:

- Informação Inicial:

$$(\theta_0 | D_0) \sim \mathcal{N}(m_0, C_0). \quad (3)$$



- Se $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}$, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:
 - Informação Inicial:

$$(\theta_0 | D_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0). \quad (3)$$



- Se $\{\mathbf{F}, \mathbf{G}, V, \mathbf{W}\}$, então o modelo é chamado de constante.
- O modelo linear dinâmico constante engloba essencialmente todos os modelos lineares tradicionais de séries temporais.
- Modelos mais gerais podem ser propostos com ν_t e ω_t correlacionados. Porém esses novos modelos podem sempre ser reescritos satisfazendo as condições de independência, West e Harrison (1997).
- A informação inicial sobre θ_t é representada por:
 - **Informação Inicial:**

$$(\theta_0|D_0) \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0). \quad (3)$$



O Teorema de Bayes

- A probabilidade de ocorrência conjunta de dois eventos distintos é dada pela regra do produto:

$$P(A,B) = P(A|B)P(B). \quad (4)$$

- Teorema de Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad (5)$$

- No contexto dos modelos lineares dinâmicos,

$$f(\theta_t|D_t) = f(\theta_t|y_t, D_{t-1}) = \frac{g(y_t|\theta_t, D_{t-1})\pi(\theta_t|D_{t-1})}{h(y_t|D_{t-1})}. \quad (6)$$



O Teorema de Bayes

- A probabilidade de ocorrência conjunta de dois eventos distintos é dada pela regra do produto:

$$P(A,B) = P(A|B)P(B). \quad (4)$$

- Teorema de Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad (5)$$

- No contexto dos modelos lineares dinâmicos,

$$f(\theta_t|D_t) = f(\theta_t|y_t, D_{t-1}) = \frac{g(y_t|\theta_t, D_{t-1})\pi(\theta_t|D_{t-1})}{h(y_t|D_{t-1})}. \quad (6)$$



O Teorema de Bayes

- A probabilidade de ocorrência conjunta de dois eventos distintos é dada pela regra do produto:

$$P(A,B) = P(A|B)P(B). \quad (4)$$

- Teorema de Bayes:

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)}. \quad (5)$$

- No contexto dos modelos lineares dinâmicos,

$$f(\theta_t|D_t) = f(\theta_t|y_t, D_{t-1}) = \frac{g(y_t|\theta_t, D_{t-1})\pi(\theta_t|D_{t-1})}{h(y_t|D_{t-1})}. \quad (6)$$



Equações de Atualização

- Seja, para algum \mathbf{m}_t e \mathbf{C}_t , a *posteriori* de θ_t :

$$(\theta_t | D_t) \sim \mathcal{N}(\mathbf{m}_t, \mathbf{C}_t). \quad (7)$$

- As equações (2) e (7) permitem calcular a distribuição *a priori* $(\theta_{t+1} | D_t)$:

$$\begin{aligned} \theta_{t+1} &= \mathbf{G}\theta_t + \omega_{t+1}, \\ (\theta_{t+1} | D_t) &\sim \mathbf{G}\mathcal{N}_{(\theta_t | D_t)}(\mathbf{m}_t, \mathbf{C}_t) + \mathcal{N}_{\omega_{t+1}}(\mathbf{0}, \mathbf{W}), \\ (\theta_{t+1} | D_t) &\sim \mathcal{N}(\mathbf{a}_{t+1}, \mathbf{R}_{t+1}), \end{aligned} \quad (8)$$

onde $\mathbf{a}_{t+1} = \mathbf{G}\mathbf{m}_t$ e $\mathbf{R}_{t+1} = \mathbf{G}\mathbf{C}_t\mathbf{G}' + \mathbf{W}$.



Equações de Atualização

- Seja, para algum \mathbf{m}_t e \mathbf{C}_t , a *posteriori* de $\boldsymbol{\theta}_t$:

$$(\boldsymbol{\theta}_t | D_t) \sim \mathcal{N}(\mathbf{m}_t, \mathbf{C}_t). \quad (7)$$

- As equações (2) e (7) permitem calcular a distribuição *a priori* $(\boldsymbol{\theta}_{t+1} | D_t)$:

$$\begin{aligned} \boldsymbol{\theta}_{t+1} &= \mathbf{G}\boldsymbol{\theta}_t + \boldsymbol{\omega}_{t+1}, \\ (\boldsymbol{\theta}_{t+1} | D_t) &\sim \mathbf{G}\mathcal{N}_{(\boldsymbol{\theta}_t | D_t)}(\mathbf{m}_t, \mathbf{C}_t) + \mathcal{N}_{\boldsymbol{\omega}_{t+1}}(\mathbf{0}, \mathbf{W}), \\ (\boldsymbol{\theta}_{t+1} | D_t) &\sim \mathcal{N}(\mathbf{a}_{t+1}, \mathbf{R}_{t+1}), \end{aligned} \quad (8)$$

onde $\mathbf{a}_{t+1} = \mathbf{G}\mathbf{m}_t$ e $\mathbf{R}_{t+1} = \mathbf{G}\mathbf{C}_t\mathbf{G}' + \mathbf{W}$.



- A equação (1) é utilizada para calcular a distribuição da previsão da observação ($Y_{t+1}|D_t$):

$$\begin{aligned} Y_{t+1} &= \mathbf{F}'\boldsymbol{\theta}_{t+1} + \nu_{t+1}, \\ (Y_{t+1}|D_t) &\sim \mathbf{F}'\mathcal{N}(\boldsymbol{\theta}_{t+1}|D_t)(\mathbf{a}_{t+1}, \mathbf{R}_{t+1}) + \mathcal{N}_{\nu_{t+1}}(0, V), \\ (Y_{t+1}|D_t) &\sim \mathcal{N}(f_{t+1}, Q_{t+1}), \end{aligned} \tag{9}$$

onde $f_{t+1} = \mathbf{F}'\mathbf{a}_{t+1}$ e $Q_{t+1} = \mathbf{F}'\mathbf{R}_{t+1}\mathbf{F} + V$.



- A distribuição *a posteriori* ($\theta_{t+1}|D_{t+1}$) é obtida via o Teorema de Bayes:

$$f(\theta_{t+1}|D_{t+1}) \propto g(Y_{t+1}|\theta_{t+1}, D_t)\pi(\theta_{t+1}|D_t) \quad (10)$$

- Desse modo:

$$(\theta_{t+1}|D_{t+1}) \sim \mathcal{N}(\mathbf{m}_{t+1}, \mathbf{C}_{t+1}), \quad (11)$$

onde, $\mathbf{m}_{t+1} = \mathbf{a}_{t+1} + \mathbf{A}_{t+1}\mathbf{e}_{t+1}$ e

$\mathbf{C}_{t+1} = \mathbf{R}_{t+1} - \mathbf{A}_{t+1}\mathbf{Q}_{t+1}\mathbf{A}_{t+1}'$, com $\mathbf{A}_{t+1} = \mathbf{R}_{t+1}\mathbf{F}\mathbf{Q}_{t+1}^{-1}$ e $\mathbf{e}_{t+1} = Y_{t+1} - f_{t+1}$, como mostra West e Harrison (1997).



- A distribuição *a posteriori* ($\theta_{t+1}|D_{t+1}$) é obtida via o Teorema de Bayes:

$$f(\theta_{t+1}|D_{t+1}) \propto g(Y_{t+1}|\theta_{t+1}, D_t)\pi(\theta_{t+1}|D_t) \quad (10)$$

- Desse modo:

$$(\theta_{t+1}|D_{t+1}) \sim \mathcal{N}(\mathbf{m}_{t+1}, \mathbf{C}_{t+1}), \quad (11)$$

onde, $\mathbf{m}_{t+1} = \mathbf{a}_{t+1} + \mathbf{A}_{t+1}\mathbf{e}_{t+1}$ e

$\mathbf{C}_{t+1} = \mathbf{R}_{t+1} - \mathbf{A}_{t+1}\mathbf{Q}_{t+1}\mathbf{A}_{t+1}'$, com $\mathbf{A}_{t+1} = \mathbf{R}_{t+1}\mathbf{F}\mathbf{Q}_{t+1}^{-1}$ e $\mathbf{e}_{t+1} = Y_{t+1} - f_{t+1}$, como mostra West e Harrison (1997).



Modelo de Volatilidade Estocástica (MVE)

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) define o modelo canônico como:
-

$$\text{Modelo Canônico : } \begin{cases} Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_0 &\sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right). \end{cases} \quad (12)$$



Modelo de Volatilidade Estocástica (MVE)

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) define o modelo canônico como:

■

$$\text{Modelo Canônico : } \begin{cases} Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_0 &\sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right). \end{cases} \quad (12)$$



Modelo de Volatilidade Estocástica (MVE)

- Modelagem inicialmente proposta por Taylor (1982).
- Kim et al. (1998) define o modelo canônico como:
-

$$\text{Modelo Canônico : } \begin{cases} Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_0 &\sim \mathcal{N}\left(\mu, \frac{\sigma_\eta^2}{1-\phi^2}\right). \end{cases} \quad (12)$$



- Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0, 1) \quad \text{e} \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \quad (13)$$

- Observe que $(Y_t | h_t) \sim \mathcal{N}(0, e^{h_t})$.
- O conjunto de parâmetros do modelo a serem estimados é definido pelo vetor:

$$\psi = (\mu, \phi, \sigma_\eta^2). \quad (14)$$



- Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0, 1) \quad \text{e} \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \quad (13)$$

- Observe que $(Y_t | h_t) \sim \mathcal{N}(0, e^{h_t})$.
- O conjunto de parâmetros do modelo a serem estimados é definido pelo vetor:

$$\psi = (\mu, \phi, \sigma_\eta^2). \quad (14)$$



- Os termos δ_t e η_t são ruídos gaussianos, independentes no tempo e entre si, com distribuições:

$$\delta_t \sim \mathcal{N}(0, 1) \quad \text{e} \quad \eta_t \sim \mathcal{N}(0, \sigma_\eta^2). \quad (13)$$

- Observe que $(Y_t|h_t) \sim \mathcal{N}(0, e^{h_t})$.
- O conjunto de parâmetros do modelo a serem estimados é definido pelo vetor:

$$\psi = (\mu, \phi, \sigma_\eta^2). \quad (14)$$



Um Modelo de Espaço-Estado

- Os MVE podem ser reescritos na notação de espaço-estado apresentada por West e Harrison (1997) estendendo-se a proposta de Zivot e Yollin (2012).
- Pela definição em (12):

$$\begin{aligned}Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ \ln Y_t^2 &= h_t + \ln \delta_t^2, \\ \ln Y_t^2 &= h_t + E(\ln \delta_t^2) + \varepsilon_t.\end{aligned}\tag{15}$$

Assim $\varepsilon_t \sim (0, \text{Var}(\ln \delta_t^2))$ é um ruído em torno de zero.



Um Modelo de Espaço-Estado

- Os MVE podem ser reescritos na notação de espaço-estado apresentada por West e Harrison (1997) estendendo-se a proposta de Zivot e Yollin (2012).
- Pela definição em (12):

$$\begin{aligned}Y_t &= e^{\frac{h_t}{2}} \delta_t, \\ \ln Y_t^2 &= h_t + \ln \delta_t^2, \\ \ln Y_t^2 &= h_t + E(\ln \delta_t^2) + \varepsilon_t.\end{aligned}\tag{15}$$

Assim $\varepsilon_t \sim (0, \text{Var}(\ln \delta_t^2))$ é um ruído em torno de zero.



- Durbin e Koopman (2012) afirmam que isso não impede a utilização das técnicas associadas aos modelos gaussianos.
- Omori et al. (2007) sugere uma mistura de $r = 10$ normais que se aproxima da distribuição do erro:

$$\ln \delta_t^2 \approx \xi_t \sim \sum_{j=1}^{10} w_j \mathcal{N}(m_j, s_j^2). \quad (16)$$



- Durbin e Koopman (2012) afirmam que isso não impede a utilização das técnicas associadas aos modelos gaussianos.
- Omori et al. (2007) sugere uma mistura de $r = 10$ normais que se aproxima da distribuição do erro:

$$\ln \delta_t^2 \approx \xi_t \sim \sum_{j=1}^{10} w_j \mathcal{N}(m_j, s_j^2). \quad (16)$$



- Comparação entre as distribuições de $\ln \delta_t^2$, que é o verdadeiro erro do modelo, de ε_t , que corresponde à aproximação ingênua, e de ξ_t , que é a mistura de 10 normais sugerida.

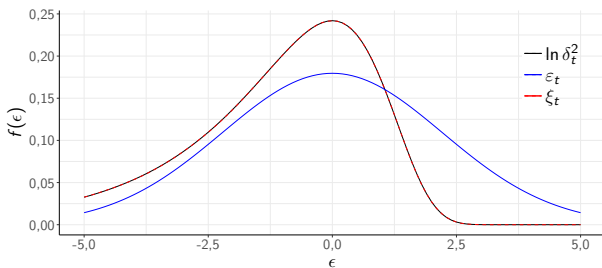


Figura: Comparação das distribuições de $\ln \delta_t^2$, ε_t e ξ_t .



- A equação que define a evolução da variável latente, h_t , em (12) pode ser reescrita:

$$\begin{aligned}h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\h_t &= (1 - \phi)\mu + \phi h_{t-1} + \eta_t.\end{aligned}\tag{17}$$

- Portanto (15) e (17) definem a equação das observações e a equação do sistema, respectivamente.



- A equação que define a evolução da variável latente, h_t , em (12) pode ser reescrita:

$$\begin{aligned}h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\h_t &= (1 - \phi)\mu + \phi h_{t-1} + \eta_t.\end{aligned}\tag{17}$$

- Portanto (15) e (17) definem a equação das observações e a equação do sistema, respectivamente.



- Assim, a equação das observações do MVE será:

$$Z_t = \ln Y_t^2 = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_t \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} + \varepsilon_t = \mathbf{F}'\boldsymbol{\theta}_t + \nu_t. \quad (18)$$

- A equação do sistema, por sua vez, será:

$$\boldsymbol{\theta}_t = \begin{bmatrix} h_t \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} = \begin{bmatrix} \phi & 1 - \phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{t-1} \\ \mu \\ E(\ln \delta_{t-1}^2) \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \\ 0 \end{bmatrix} = \mathbf{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t. \quad (19)$$



- Assim, a equação das observações do MVE será:

$$Z_t = \ln Y_t^2 = \begin{bmatrix} 1 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_t \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} + \varepsilon_t = \mathbf{F}'\boldsymbol{\theta}_t + \nu_t. \quad (18)$$

- A equação do sistema, por sua vez, será:

$$\boldsymbol{\theta}_t = \begin{bmatrix} h_t \\ \mu \\ E(\ln \delta_t^2) \end{bmatrix} = \begin{bmatrix} \phi & 1 - \phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} h_{t-1} \\ \mu \\ E(\ln \delta_{t-1}^2) \end{bmatrix} + \begin{bmatrix} \eta_t \\ 0 \\ 0 \end{bmatrix} = \mathbf{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t. \quad (19)$$



- O modelo de volatilidade estocástica é definido como um modelo de espaço-estado constante por:

$$\text{Modelo de Espaço-Estado : } \begin{cases} \ln Y_t^2 = \mathbf{F}'\boldsymbol{\theta}_t + \varepsilon_t, \\ \boldsymbol{\theta}_t = \mathbf{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \\ \boldsymbol{\theta}_0 \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0), \end{cases} \quad (20)$$

- A quádrupla que caracteriza o modelo é:

$$\{\mathbf{F}, \mathbf{G}, \mathbf{V}, \mathbf{W}\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \phi & 1-\phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{Var}(\varepsilon_t), \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\}. \quad (21)$$



- O modelo de volatilidade estocástica é definido como um modelo de espaço-estado constante por:

$$\text{Modelo de Espaço-Estado : } \begin{cases} \ln Y_t^2 = \mathbf{F}'\boldsymbol{\theta}_t + \varepsilon_t, \\ \boldsymbol{\theta}_t = \mathbf{G}\boldsymbol{\theta}_{t-1} + \boldsymbol{\omega}_t, \\ \boldsymbol{\theta}_0 \sim \mathcal{N}(\mathbf{m}_0, \mathbf{C}_0), \end{cases} \quad (20)$$

- A quádrupla que caracteriza o modelo é:

$$\{\mathbf{F}, \mathbf{G}, \mathbf{V}, \mathbf{W}\} = \left\{ \begin{bmatrix} 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} \phi & 1-\phi & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}, \text{Var}(\varepsilon_t), \begin{bmatrix} \sigma_\eta^2 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{bmatrix} \right\}. \quad (21)$$



Modelo Não Centralizado

- Kastner e Frühwirth-Schnatter (2014) propõe uma transformação linear na variável latente, h_t :

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma_\eta}. \quad (22)$$

- É possível, então, reescrever a equação do sistema, que passa a ser:

$$\begin{aligned} h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_t - \mu &= \phi(h_{t-1} - \mu) + \eta_t, \\ \left(\frac{h_t - \mu}{\sigma_\eta} \right) &= \phi \left(\frac{h_{t-1} - \mu}{\sigma_\eta} \right) + \frac{\eta_t}{\sigma_\eta}, \\ \tilde{h}_t &= \phi \tilde{h}_{t-1} + \tilde{\eta}_t. \end{aligned} \quad (23)$$



Modelo Não Centralizado

- Kastner e Frühwirth-Schnatter (2014) propõe uma transformação linear na variável latente, h_t :

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma_\eta}. \quad (22)$$

- É possível, então, reescrever a equação do sistema, que passa a ser:

$$\begin{aligned} h_t &= \mu + \phi(h_{t-1} - \mu) + \eta_t, \\ h_t - \mu &= \phi(h_{t-1} - \mu) + \eta_t, \\ \left(\frac{h_t - \mu}{\sigma_\eta} \right) &= \phi \left(\frac{h_{t-1} - \mu}{\sigma_\eta} \right) + \frac{\eta_t}{\sigma_\eta}, \\ \tilde{h}_t &= \phi \tilde{h}_{t-1} + \tilde{\eta}_t. \end{aligned} \quad (23)$$



- De modo semelhante, é possível reescrever a equação do sistema e a distribuição inicial da variável latente, o que leva ao modelo:

$$\begin{array}{l} \textbf{Modelo} \\ \textbf{Não Centralizado :} \end{array} \left\{ \begin{array}{l} Y_t = e^{\frac{\mu}{2}} e^{\sigma_\eta \frac{\tilde{h}_t}{2}} \delta_t, \\ \tilde{h}_t = \phi \tilde{h}_{t-1} + \tilde{\eta}_t, \\ \tilde{h}_0 \sim \mathcal{N}\left(0, \frac{1}{1-\phi^2}\right). \end{array} \right. \quad (24)$$



Proposta de Estimação dos Parâmetros do MVE

- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, e basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, e muitas maneiras de se estimar os parâmetros.

Tabela: Métodos de estimação dos parâmetros do MVE e principais referências, (Bos, 2012).

Método	Referência	Paradigma
Quasi-Maximum Likelihood (QML)	Harvey et al. (1994)	Clássico
Gaussian Mixture Sampling (GMS)	Kim et al. (1998)	Bayesiano
Simulated Method of Moments (SMM)	Gallant e Tauchen (1996)	Clássico
■ Importance Sampling (IS)	Durbin e Koopman (1997)	Clássico
Efficient Importance Sampling (EIS)	Richard e Zhang (2007)	Clássico
Improved Importance Sampling (IIS)	Nguyen (2007)	Clássico
Single Site Sampler (SSS)	Carter e Kohn (1994)	Bayesiano
MultiMove Sampler (MMS)	Shephard e Pitt (1997)	Bayesiano



Proposta de Estimação dos Parâmetros do MVE

- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, e basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, e muitas maneiras de se estimar os parâmetros.

Tabela: Métodos de estimação dos parâmetros do MVE e principais referências, (Bos, 2012).

Método	Referência	Paradigma
Quasi-Maximum Likelihood (QML)	Harvey et al. (1994)	Clássico
Gaussian Mixture Sampling (GMS)	Kim et al. (1998)	Bayesiano
Simulated Method of Moments (SMM)	Gallant e Tauchen (1996)	Clássico
■ Importance Sampling (IS)	Durbin e Koopman (1997)	Clássico
Efficient Importance Sampling (EIS)	Richard e Zhang (2007)	Clássico
Improved Importance Sampling (IIS)	Nguyen (2007)	Clássico
Single Site Sampler (SSS)	Carter e Kohn (1994)	Bayesiano
MultiMove Sampler (MMS)	Shephard e Pitt (1997)	Bayesiano



Proposta de Estimação dos Parâmetros do MVE

- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, e basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, e muitas maneiras de se estimar os parâmetros.

Tabela: Métodos de estimação dos parâmetros do MVE e principais referências, (Bos, 2012).

Método	Referência	Paradigma
Quasi-Maximum Likelihood (QML)	Harvey et al. (1994)	Clássico
Gaussian Mixture Sampling (GMS)	Kim et al. (1998)	Bayesiano
Simulated Method of Moments (SMM)	Gallant e Tauchen (1996)	Clássico
■ Importance Sampling (IS)	Durbin e Koopman (1997)	Clássico
Efficient Importance Sampling (EIS)	Richard e Zhang (2007)	Clássico
Improved Importance Sampling (IIS)	Nguyen (2007)	Clássico
Single Site Sampler (SSS)	Carter e Kohn (1994)	Bayesiano
MultiMove Sampler (MMS)	Shephard e Pitt (1997)	Bayesiano



Proposta de Estimação dos Parâmetros do MVE

- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, e basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, e muitas maneiras de se estimar os parâmetros.

Tabela: Métodos de estimação dos parâmetros do MVE e principais referências, (Bos, 2012).

Método	Referência	Paradigma
Quasi-Maximum Likelihood (QML)	Harvey et al. (1994)	Clássico
Gaussian Mixture Sampling (GMS)	Kim et al. (1998)	Bayesiano
Simulated Method of Moments (SMM)	Gallant e Tauchen (1996)	Clássico
■ Importance Sampling (IS)	Durbin e Koopman (1997)	Clássico
Efficient Importance Sampling (EIS)	Richard e Zhang (2007)	Clássico
Improved Importance Sampling (IIS)	Nguyen (2007)	Clássico
Single Site Sampler (SSS)	Carter e Kohn (1994)	Bayesiano
MultiMove Sampler (MMS)	Shephard e Pitt (1997)	Bayesiano



Proposta de Estimação dos Parâmetros do MVE

- MVE vs ARCH e GARCH.
- Bos (2012) cita que o MVE é pouco amigável.
 - ARCH e GARCH: Muitas variações do modelo, e basicamente uma maneira de se estimar os parâmetros.
 - MVE: Basicamente uma definição do modelo, e muitas maneiras de se estimar os parâmetros.

Tabela: Métodos de estimação dos parâmetros do MVE e principais referências, (Bos, 2012).

Método	Referência	Paradigma
Quasi-Maximum Likelihood (QML)	Harvey et al. (1994)	Clássico
Gaussian Mixture Sampling (GMS)	Kim et al. (1998)	Bayesiano
Simulated Method of Moments (SMM)	Gallant e Tauchen (1996)	Clássico
▪ Importance Sampling (IS)	Durbin e Koopman (1997)	Clássico
Efficient Importance Sampling (EIS)	Richard e Zhang (2007)	Clássico
Improved Importance Sampling (IIS)	Nguyen (2007)	Clássico
Single Site Sampler (SSS)	Carter e Kohn (1994)	Bayesiano
MultiMove Sampler (MMS)	Shephard e Pitt (1997)	Bayesiano



Definição do Modelo Bayesiano

- A estimação Bayesiana dos parâmetros do MVE consiste em determinar a distribuição *a posteriori* conjunta de:

$$\psi = (\mu, \phi, \sigma_\eta^2). \quad (25)$$

- Porém, é razoável supor:

$$p_\psi(\psi) = p_\mu(\mu)p_\phi(\phi)p_{\sigma_\eta^2}(\sigma_\eta^2). \quad (26)$$

- Como o paradigma Bayesiano sugere, deve-se definir as distribuições *a priori* dos parâmetros.
- As distribuições *a priori* foram definidas conforme Kastner e Frühwirth-Schnatter (2014) e Kim et al. (1998).



Definição do Modelo Bayesiano

- A estimação Bayesiana dos parâmetros do MVE consiste em determinar a distribuição *a posteriori* conjunta de:

$$\psi = (\mu, \phi, \sigma_\eta^2). \quad (25)$$

- Porém, é razoável supor:

$$p_\psi(\psi) = p_\mu(\mu)p_\phi(\phi)p_{\sigma_\eta^2}(\sigma_\eta^2). \quad (26)$$

- Como o paradigma Bayesiano sugere, deve-se definir as distribuições *a priori* dos parâmetros.
- As distribuições *a priori* foram definidas conforme Kastner e Frühwirth-Schnatter (2014) e Kim et al. (1998).



Definição do Modelo Bayesiano

- A estimação Bayesiana dos parâmetros do MVE consiste em determinar a distribuição *a posteriori* conjunta de:

$$\psi = (\mu, \phi, \sigma_\eta^2). \quad (25)$$

- Porém, é razoável supor:

$$p_\psi(\psi) = p_\mu(\mu)p_\phi(\phi)p_{\sigma_\eta^2}(\sigma_\eta^2). \quad (26)$$

- Como o paradigma Bayesiano sugere, deve-se definir as distribuições *a priori* dos parâmetros.
- As distribuições *a priori* foram definidas conforme Kastner e Frühwirth-Schnatter (2014) e Kim et al. (1998).



Definição do Modelo Bayesiano

- A estimação Bayesiana dos parâmetros do MVE consiste em determinar a distribuição *a posteriori* conjunta de:

$$\psi = (\mu, \phi, \sigma_\eta^2). \quad (25)$$

- Porém, é razoável supor:

$$p_\psi(\psi) = p_\mu(\mu)p_\phi(\phi)p_{\sigma_\eta^2}(\sigma_\eta^2). \quad (26)$$

- Como o paradigma Bayesiano sugere, deve-se definir as distribuições *a priori* dos parâmetros.
- As distribuições *a priori* foram definidas conforme Kastner e Frühwirth-Schnatter (2014) e Kim et al. (1998).



- O nível μ da volatilidade tem seu suporte em \mathbb{R} , e será atribuída uma distribuição *a priori*, $\pi(\mu)$, com densidade Gaussiana para esse parâmetro:

$$\mu \sim \mathcal{N}(a_\mu, B_\mu). \quad (27)$$

- O parâmetro ϕ determina a persistência da volatilidade, e $|\phi| < 1$. Seja uma nova variável ϕ_0 cuja distribuição *a priori* é $\mathcal{B}(a_\phi, b_\phi)$.
- A persistência então será $\phi = 2\phi_0 - 1$ e sua distribuição *a priori* será:

$$\pi(\phi) = \frac{\Gamma(a_\phi + b_\phi)}{2\Gamma(a_\phi)\Gamma(b_\phi)} \left(\frac{1+\phi}{2}\right)^{a_\phi-1} \left(\frac{1-\phi}{2}\right)^{b_\phi-1}. \quad (28)$$



- O nível μ da volatilidade tem seu suporte em \mathbb{R} , e será atribuída uma distribuição *a priori*, $\pi(\mu)$, com densidade Gaussiana para esse parâmetro:

$$\mu \sim \mathcal{N}(a_\mu, B_\mu). \quad (27)$$

- O parâmetro ϕ determina a persistência da volatilidade, e $|\phi| < 1$. Seja uma nova variável ϕ_0 cuja distribuição *a priori* é $\mathcal{B}(a_\phi, b_\phi)$.
- A persistência então será $\phi = 2\phi_0 - 1$ e sua distribuição *a priori* será:

$$\pi(\phi) = \frac{\Gamma(a_\phi + b_\phi)}{2\Gamma(a_\phi)\Gamma(b_\phi)} \left(\frac{1+\phi}{2}\right)^{a_\phi-1} \left(\frac{1-\phi}{2}\right)^{b_\phi-1}. \quad (28)$$



- O nível μ da volatilidade tem seu suporte em \mathbb{R} , e será atribuída uma distribuição *a priori*, $\pi(\mu)$, com densidade Gaussiana para esse parâmetro:

$$\mu \sim \mathcal{N}(a_\mu, B_\mu). \quad (27)$$

- O parâmetro ϕ determina a persistência da volatilidade, e $|\phi| < 1$. Seja uma nova variável ϕ_0 cuja distribuição *a priori* é $\mathcal{B}(a_\phi, b_\phi)$.
- A persistência então será $\phi = 2\phi_0 - 1$ e sua distribuição *a priori* será:

$$\pi(\phi) = \frac{\Gamma(a_\phi + b_\phi)}{2\Gamma(a_\phi)\Gamma(b_\phi)} \left(\frac{1+\phi}{2}\right)^{a_\phi-1} \left(\frac{1-\phi}{2}\right)^{b_\phi-1}. \quad (28)$$



- A variância da volatilidade, σ_η^2 , tem seus possíveis valores em \mathbb{R}^+ . Sua distribuição *a priori* será:

$$\sigma_\eta^2 \sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2B_\sigma}\right). \quad (29)$$



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



Estimando μ , ϕ e σ_η^2

- A estimação dos parâmetros será feita via MCMC.
- Kastner e Frühwirth-Schnatter (2014) propõe uma estratégia chamada de *Ancillarity-Sufficiency Interweaving Strategy* (ASIS).
- Algoritmo da ASIS:
 1. Iniciar os valores de μ , ϕ e σ_η^2 ;
 2. Estimar a série de h_t (modelo canônico);
 3. A partir dos valores de h_t , estimar os valores de μ , ϕ e σ_η^2 ;
 4. Transformar a série de h_t para o modelo não centralizado, h_t^* ;
 5. A partir dos valores de h_t^* , estimar novamente os valores de μ , ϕ e σ_η^2 .



- Baseado em Yu e Meng (2011), Kastner e Frühwirth-Schnatter (2014) diz que a variável latente, no modelo canônico, forma uma estatística suficiente para μ e σ_{η}^2 . E quando h_t é transformada para o modelo não centralizado, forma uma estatística ancilar.
- Alternar entre essas especificações do modelo aumenta a eficiência do amostrador via MCMC.



- Baseado em Yu e Meng (2011), Kastner e Frühwirth-Schnatter (2014) diz que a variável latente, no modelo canônico, forma uma estatística suficiente para μ e σ_{η}^2 . E quando h_t é transformada para o modelo não centralizado, forma uma estatística ancilar.
- Alternar entre essas especificações do modelo aumenta a eficiência do amostrador via MCMC.



- Sob o modelo canônico, os parâmetros são estimados pelo algoritmo de Metropolis-Hastings (Metropolis et al. (1953); Hastings (1970)).
- Geradora de candidatos:
 - σ_{η}^2 : Distribuição Gamma-Inversa;
 - ϕ : Distribuição Normal truncada;
 - μ : Distribuição Normal (amostrado indiretamente através da variável $\gamma = (1 - \phi)\mu$).
- O candidato é aceito com probabilidade $\min(1, R)$.



- Sob o modelo canônico, os parâmetros são estimados pelo algoritmo de Metropolis-Hastings (Metropolis et al. (1953); Hastings (1970)).
- Geradora de candidatos:
 - σ_η^2 : Distribuição Gamma-Inversa;
 - ϕ : Distribuição Normal truncada;
 - μ : Distribuição Normal (amostrado indiretamente através da variável $\gamma = (1 - \phi)\mu$).
- O candidato é aceito com probabilidade $\min(1, R)$.



- Sob o modelo canônico, os parâmetros são estimados pelo algoritmo de Metropolis-Hastings (Metropolis et al. (1953); Hastings (1970)).
- Geradora de candidatos:
 - σ_{η}^2 : Distribuição Gamma-Inversa;
 - ϕ : Distribuição Normal truncada;
 - μ : Distribuição Normal (amostrado indiretamente através da variável $\gamma = (1 - \phi)\mu$).
- O candidato é aceito com probabilidade $\min(1, R)$.



- Sob o modelo canônico, os parâmetros são estimados pelo algoritmo de Metropolis-Hastings (Metropolis et al. (1953); Hastings (1970)).
- Geradora de candidatos:
 - σ_η^2 : Distribuição Gamma-Inversa;
 - ϕ : Distribuição Normal truncada;
 - μ : Distribuição Normal (amostrado indiretamente através da variável $\gamma = (1 - \phi)\mu$).
- O candidato é aceito com probabilidade $\min(1, R)$.



- Sob o modelo canônico, os parâmetros são estimados pelo algoritmo de Metropolis-Hastings (Metropolis et al. (1953); Hastings (1970)).
- Geradora de candidatos:
 - σ_η^2 : Distribuição Gamma-Inversa;
 - ϕ : Distribuição Normal truncada;
 - μ : Distribuição Normal (amostrado indiretamente através da variável $\gamma = (1 - \phi)\mu$).
- O candidato é aceito com probabilidade $\min(1, R)$.



- Sob o modelo canônico, os parâmetros são estimados pelo algoritmo de Metropolis-Hastings (Metropolis et al. (1953); Hastings (1970)).
- Geradora de candidatos:
 - σ_η^2 : Distribuição Gamma-Inversa;
 - ϕ : Distribuição Normal truncada;
 - μ : Distribuição Normal (amostrado indiretamente através da variável $\gamma = (1 - \phi)\mu$).
- O candidato é aceito com probabilidade $\min(1, R)$.



- A variável latente é transformada para o modelo não centralizado via:

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma_\eta}. \quad (30)$$

- Os valores de μ e σ_η^2 são reamostrados, agora diretamente, através das novas geradoras de candidatos:
 - σ_η : Distribuição Normal;
 - μ : Distribuição Normal.
- O valor de ϕ não precisa ser reamostrado, pois esse parâmetro não é explicitamente envolvido na transformação da variável latente.



- A variável latente é transformada para o modelo não centralizado via:

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma_\eta}. \quad (30)$$

- Os valores de μ e σ_η^2 são reamostrados, agora diretamente, através das novas geradoras de candidatos:
 - σ_η : Distribuição Normal;
 - μ : Distribuição Normal.
- O valor de ϕ não precisa ser reamostrado, pois esse parâmetro não é explicitamente envolvido na transformação da variável latente.



- A variável latente é transformada para o modelo não centralizado via:

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma_\eta}. \quad (30)$$

- Os valores de μ e σ_η^2 são reamostrados, agora diretamente, através das novas geradoras de candidatos:
 - σ_η : Distribuição Normal;
 - μ : Distribuição Normal.
- O valor de ϕ não precisa ser reamostrado, pois esse parâmetro não é explicitamente envolvido na transformação da variável latente.



- A variável latente é transformada para o modelo não centralizado via:

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma_\eta}. \quad (30)$$

- Os valores de μ e σ_η^2 são reamostrados, agora diretamente, através das novas geradoras de candidatos:
 - σ_η : Distribuição Normal;
 - μ : Distribuição Normal.
- O valor de ϕ não precisa ser reamostrado, pois esse parâmetro não é explicitamente envolvido na transformação da variável latente.



- A variável latente é transformada para o modelo não centralizado via:

$$\tilde{h}_t = \frac{h_t - \mu}{\sigma_\eta}. \quad (30)$$

- Os valores de μ e σ_η^2 são reamostrados, agora diretamente, através das novas geradoras de candidatos:
 - σ_η : Distribuição Normal;
 - μ : Distribuição Normal.
- O valor de ϕ não precisa ser reamostrado, pois esse parâmetro não é explicitamente envolvido na transformação da variável latente.



Estimando $h_{1:N}$

- Kastner e Frühwirth-Schnatter (2014) apresenta uma maneira de se estimar a variável latente através de uma distribuição normal N —variada.
- A implementação computacional desse algoritmo não é trivial.
- Proposta alternativa: estimar h_t através de um algoritmo inspirado em McCormick et al. (2012).
- Essa proposta é motivada no fato de ser mais simples do que a proposta original em Kastner e Frühwirth-Schnatter (2014).



Estimando $h_{1:N}$

- Kastner e Frühwirth-Schnatter (2014) apresenta uma maneira de se estimar a variável latente através de uma distribuição normal N –variada.
- A implementação computacional desse algoritmo não é trivial.
- Proposta alternativa: estimar h_t através de um algoritmo inspirado em McCormick et al. (2012).
- Essa proposta é motivada no fato de ser mais simples do que a proposta original em Kastner e Frühwirth-Schnatter (2014).



Estimando $h_{1:N}$

- Kastner e Frühwirth-Schnatter (2014) apresenta uma maneira de se estimar a variável latente através de uma distribuição normal N —variada.
- A implementação computacional desse algoritmo não é trivial.
- Proposta alternativa: estimar h_t através de um algoritmo inspirado em McCormick et al. (2012).
- Essa proposta é motivada no fato de ser mais simples do que a proposta original em Kastner e Frühwirth-Schnatter (2014).



Estimando $h_{1:N}$

- Kastner e Frühwirth-Schnatter (2014) apresenta uma maneira de se estimar a variável latente através de uma distribuição normal N —variada.
- A implementação computacional desse algoritmo não é trivial.
- Proposta alternativa: estimar h_t através de um algoritmo inspirado em McCormick et al. (2012).
- Essa proposta é motivada no fato de ser mais simples do que a proposta original em Kastner e Frühwirth-Schnatter (2014).



O Método de McCormick et al. (2012)

- A estimação recursiva dos estados latentes, θ_t , do modelo começa supondo que:

$$(\theta_{t-1}|D_{t-1}) \sim \mathcal{N}(\hat{\theta}_{t-1}, \hat{\Sigma}_{t-1}). \quad (31)$$

- A equação de predição é então descrita por:

$$(\theta_t|D_{t-1}) \sim \mathcal{N}(\mathbf{G}\hat{\theta}_{t-1}, \mathbf{R}_t), \quad (32)$$

em que

$$\mathbf{R}_t = \frac{\mathbf{G}\hat{\Sigma}_{t-1}\mathbf{G}^T}{\lambda_t}, \quad (33)$$

- λ_t é um fator de desconto com valores no intervalo $(0, 1)$.



O Método de McCormick et al. (2012)

- A estimação recursiva dos estados latentes, θ_t , do modelo começa supondo que:

$$(\theta_{t-1}|D_{t-1}) \sim \mathcal{N}(\hat{\theta}_{t-1}, \hat{\Sigma}_{t-1}). \quad (31)$$

- A equação de predição é então descrita por:

$$(\theta_t|D_{t-1}) \sim \mathcal{N}(\mathbf{G}\hat{\theta}_{t-1}, \mathbf{R}_t), \quad (32)$$

em que

$$\mathbf{R}_t = \frac{\mathbf{G}\hat{\Sigma}_{t-1}\mathbf{G}^T}{\lambda_t}, \quad (33)$$

- λ_t é um fator de desconto com valores no intervalo $(0, 1)$.



O Método de McCormick et al. (2012)

- A estimação recursiva dos estados latentes, θ_t , do modelo começa supondo que:

$$(\theta_{t-1}|D_{t-1}) \sim \mathcal{N}(\hat{\theta}_{t-1}, \hat{\Sigma}_{t-1}). \quad (31)$$

- A equação de predição é então descrita por:

$$(\theta_t|D_{t-1}) \sim \mathcal{N}(\mathbf{G}\hat{\theta}_{t-1}, \mathbf{R}_t), \quad (32)$$

em que

$$\mathbf{R}_t = \frac{\mathbf{G}\hat{\Sigma}_{t-1}\mathbf{G}^T}{\lambda_t}, \quad (33)$$

- λ_t é um fator de desconto com valores no intervalo $(0, 1)$.



- A distribuição *a posteriori* de $(\theta_t|D_t)$ é obtida aproximadamente usando-se (32) e o Teorema de Bayes:

$$p(\theta_t|D_t) \propto p(y_t|\theta_t)\mathcal{N}(\mathbf{G}\hat{\theta}_{t-1}, \mathbf{R}_t). \quad (34)$$

- O lado direito de (34) não possui forma fechada, então a estimativa de θ_t é dada por:

$$\hat{\theta}_t = \hat{\theta}_{t-1} - \left[D^2 l(\hat{\theta}_{t-1}) \right]^{-1} D l(\hat{\theta}_{t-1}), \quad (35)$$

onde $Dl(\theta_t)$ e $D^2l(\theta_t)$ são a primeira e a segunda derivadas de $l(\theta_t) = \ln p(\theta_t|D_t)$.



- A distribuição *a posteriori* de $(\theta_t|D_t)$ é obtida aproximadamente usando-se (32) e o Teorema de Bayes:

$$p(\theta_t|D_t) \propto p(y_t|\theta_t)\mathcal{N}(\mathbf{G}\hat{\theta}_{t-1}, \mathbf{R}_t). \quad (34)$$

- O lado direito de (34) não possui forma fechada, então a estimativa de θ_t é dada por:

$$\hat{\theta}_t = \hat{\theta}_{t-1} - \left[D^2 l(\hat{\theta}_{t-1}) \right]^{-1} D l(\hat{\theta}_{t-1}), \quad (35)$$

onde $Dl(\theta_t)$ e $D^2l(\theta_t)$ são a primeira e a segunda derivadas de $l(\theta_t) = \ln p(\theta_t|D_t)$.



- Para atualizar a matriz de covariância Σ_t , utiliza-se:

$$\hat{\Sigma}_t = - \left[D^2 l(\hat{\theta}_{t-1}) \right]^{-1}. \quad (36)$$

- O valor de λ_t é obtido de forma a maximizar a distribuição preditiva:

$$p(Y_t | D_{t-1}, \lambda_t) = \int p(Y_t | \theta_t, D_{t-1}) p(\theta_t | D_{t-1}, \lambda_t) d\theta_t. \quad (37)$$

- Os autores utilizam aproximações de Laplace (Tierney e Kadane, 1986) para calcular o valor de (37).



- Para atualizar a matriz de covariância Σ_t , utiliza-se:

$$\hat{\Sigma}_t = - \left[D^2 l(\hat{\theta}_{t-1}) \right]^{-1}. \quad (36)$$

- O valor de λ_t é obtido de forma a maximizar a distribuição preditiva:

$$p(Y_t | D_{t-1}, \lambda_t) = \int p(Y_t | \theta_t, D_{t-1}) p(\theta_t | D_{t-1}, \lambda_t) d\theta_t. \quad (37)$$

- Os autores utilizam aproximações de Laplace (Tierney e Kadane, 1986) para calcular o valor de (37).



- Para atualizar a matriz de covariância Σ_t , utiliza-se:

$$\hat{\Sigma}_t = - \left[D^2 l(\hat{\theta}_{t-1}) \right]^{-1}. \quad (36)$$

- O valor de λ_t é obtido de forma a maximizar a distribuição preditiva:

$$p(Y_t | D_{t-1}, \lambda_t) = \int p(Y_t | \theta_t, D_{t-1}) p(\theta_t | D_{t-1}, \lambda_t) d\theta_t. \quad (37)$$

- Os autores utilizam aproximações de Laplace (Tierney e Kadane, 1986) para calcular o valor de (37).



Adaptação do Método de McCormick et al. (2012) ao MVE

- Seja, por hipótese:

$$(h_{t-1}|D_{t-1}) \sim \mathcal{N}(\hat{m}_{t-1}, \hat{C}_{t-1}). \quad (38)$$

- A distribuição de predição $(h_t|D_{t-1})$ é encontrada através da equação do sistema e de (38):

$$(h_t|D_{t-1}) \sim \mathcal{N}(\mu(1 - \phi) + \phi\hat{m}_{t-1}, \phi^2\hat{C}_{t-1} + \sigma_\eta^2). \quad (39)$$

- Nesse caso:

$$R_t = \frac{\phi^2\hat{C}_{t-1}}{\lambda_t}. \quad (40)$$



Adaptação do Método de McCormick et al. (2012) ao MVE

- Seja, por hipótese:

$$(h_{t-1}|D_{t-1}) \sim \mathcal{N}(\hat{m}_{t-1}, \hat{C}_{t-1}). \quad (38)$$

- A distribuição de predição $(h_t|D_{t-1})$ é encontrada através da equação do sistema e de (38):

$$(h_t|D_{t-1}) \sim \mathcal{N}(\mu(1 - \phi) + \phi\hat{m}_{t-1}, \phi^2\hat{C}_{t-1} + \sigma_\eta^2). \quad (39)$$

- Nesse caso:

$$R_t = \frac{\phi^2\hat{C}_{t-1}}{\lambda_t}. \quad (40)$$



Adaptação do Método de McCormick et al. (2012) ao MVE

- Seja, por hipótese:

$$(h_{t-1}|D_{t-1}) \sim \mathcal{N}(\hat{m}_{t-1}, \hat{C}_{t-1}). \quad (38)$$

- A distribuição de predição $(h_t|D_{t-1})$ é encontrada através da equação do sistema e de (38):

$$(h_t|D_{t-1}) \sim \mathcal{N}(\mu(1 - \phi) + \phi\hat{m}_{t-1}, \phi^2\hat{C}_{t-1} + \sigma_\eta^2). \quad (39)$$

- Nesse caso:

$$R_t = \frac{\phi^2\hat{C}_{t-1}}{\lambda_t}. \quad (40)$$



- Pelo Teorema de Bayes, a distribuição *a posteriori* ($h_t|D_t$) será:

$$p(h_t|D_t) \propto p(y_t|h_t)p(h_t|D_{t-1}). \quad (41)$$

- O logaritmo do lado direito de (41) é dado por:

$$l(h_t|\mu, \phi) \propto -\frac{h_t}{2} - \frac{y_t^2}{2e^{h_t}} - \frac{1}{2} \ln R_t - \frac{1}{2R_t} [h_t - (\mu(1 - \phi) + \phi\hat{m}_{t-1})]^2. \quad (42)$$

- Com isso, iterativamente ao longo das N observações, o estimador, \hat{h}_t , da variável latente, h_t , será:

$$\hat{h}_t = \hat{m}_t = \hat{m}_{t-1} - \frac{l'(\hat{m}_{t-1})}{l''(\hat{m}_{t-1})}. \quad (43)$$



- Pelo Teorema de Bayes, a distribuição *a posteriori* ($h_t|D_t$) será:

$$p(h_t|D_t) \propto p(y_t|h_t)p(h_t|D_{t-1}). \quad (41)$$

- O logaritmo do lado direito de (41) é dado por:

$$l(h_t|\mu, \phi) \propto -\frac{h_t}{2} - \frac{y_t^2}{2e^{h_t}} - \frac{1}{2} \ln R_t - \frac{1}{2R_t} [h_t - (\mu(1 - \phi) + \phi\hat{m}_{t-1})]^2. \quad (42)$$

- Com isso, iterativamente ao longo das N observações, o estimador, \hat{h}_t , da variável latente, h_t , será:

$$\hat{h}_t = \hat{m}_t = \hat{m}_{t-1} - \frac{l'(\hat{m}_{t-1})}{l''(\hat{m}_{t-1})}. \quad (43)$$



- Pelo Teorema de Bayes, a distribuição *a posteriori* ($h_t|D_t$) será:

$$p(h_t|D_t) \propto p(y_t|h_t)p(h_t|D_{t-1}). \quad (41)$$

- O logaritmo do lado direito de (41) é dado por:

$$l(h_t|\mu, \phi) \propto -\frac{h_t}{2} - \frac{y_t^2}{2e^{h_t}} - \frac{1}{2} \ln R_t - \frac{1}{2R_t} [h_t - (\mu(1 - \phi) + \phi\hat{m}_{t-1})]^2. \quad (42)$$

- Com isso, iterativamente ao longo das N observações, o estimador, \hat{h}_t , da variável latente, h_t , será:

$$\hat{h}_t = \hat{m}_t = \hat{m}_{t-1} - \frac{l'(\hat{m}_{t-1})}{l''(\hat{m}_{t-1})}. \quad (43)$$



- A primeira derivada é:

$$l'(h_t|\mu, \phi) = -\frac{1}{2} + \frac{y_t^2}{2e^{h_t}} - \frac{1}{R_t} [h_t - (\mu(1 - \phi) + \phi\hat{m}_{t-1})]. \quad (44)$$

- A segunda derivada, por sua vez, é:

$$l''(h_t|\mu, \phi) = -\frac{y_t^2}{2e^{h_t}} - \frac{1}{R_t}. \quad (45)$$

- A variância da distribuição de h_t ao longo do processo é atualizada por:

$$\hat{C}_t = -\frac{1}{l'''(\hat{m}_{t-1})}. \quad (46)$$



- A primeira derivada é:

$$l'(h_t|\mu, \phi) = -\frac{1}{2} + \frac{y_t^2}{2e^{h_t}} - \frac{1}{R_t} [h_t - (\mu(1 - \phi) + \phi\hat{m}_{t-1})]. \quad (44)$$

- A segunda derivada, por sua vez, é:

$$l''(h_t|\mu, \phi) = -\frac{y_t^2}{2e^{h_t}} - \frac{1}{R_t}. \quad (45)$$

- A variância da distribuição de h_t ao longo do processo é atualizada por:

$$\hat{C}_t = -\frac{1}{l'''(\hat{m}_{t-1})}. \quad (46)$$



- A primeira derivada é:

$$l'(h_t|\mu, \phi) = -\frac{1}{2} + \frac{y_t^2}{2e^{h_t}} - \frac{1}{R_t} [h_t - (\mu(1 - \phi) + \phi\hat{m}_{t-1})]. \quad (44)$$

- A segunda derivada, por sua vez, é:

$$l''(h_t|\mu, \phi) = -\frac{y_t^2}{2e^{h_t}} - \frac{1}{R_t}. \quad (45)$$

- A variância da distribuição de h_t ao longo do processo é atualizada por:

$$\hat{C}_t = -\frac{1}{l'''(\hat{m}_{t-1})}. \quad (46)$$



- O valor do fator de desconto, λ_t , é calculado de modo a maximizar a distribuição preditiva ($y_t|D_{t-1}$):

$$p(y_t|D_{t-1}, \lambda_t) = \int_{h_t} p(y_t|h_t, D_{t-1})p(h_t|D_{t-1}, \lambda_t)dh_t. \quad (47)$$

- Como a integral não possui forma fechada, o valor da distribuição preditiva é tomado por uma aproximação de Laplace.



- O valor do fator de desconto, λ_t , é calculado de modo a maximizar a distribuição preditiva ($y_t|D_{t-1}$):

$$p(y_t|D_{t-1}, \lambda_t) = \int_{h_t} p(y_t|h_t, D_{t-1})p(h_t|D_{t-1}, \lambda_t)dh_t. \quad (47)$$

- Como a integral não possui forma fechada, o valor da distribuição preditiva é tomado por uma aproximação de Laplace.



Aplicação da Metodologia Proposta

- A metodologia proposta foi aplicada num conjunto de dados simulados.
- As séries com 1.461 observações se caracterizam pelo produto cruzado dos possíveis valores predeterminados para μ , ϕ e σ_{η}^2 a seguir:
 - $\mu = -5,4$,
 - $\phi \in \{0,50; 0,80; 0,90; 0,99\}$,
 - $\sigma_{\eta}^2 \in \{0,1^2; 0,3^2; 0,5^2\}$.



Aplicação da Metodologia Proposta

- A metodologia proposta foi aplicada num conjunto de dados simulados.
- As séries com 1.461 observações se caracterizam pelo produto cruzado dos possíveis valores predeterminados para μ , ϕ e σ_{η}^2 a seguir:
 - $\mu = -5,4$,
 - $\phi \in \{0,50; 0,80; 0,90; 0,99\}$,
 - $\sigma_{\eta}^2 \in \{0,1^2; 0,3^2; 0,5^2\}$.



Aplicação da Metodologia Proposta

- A metodologia proposta foi aplicada num conjunto de dados simulados.
- As séries com 1.461 observações se caracterizam pelo produto cruzado dos possíveis valores predeterminados para μ , ϕ e σ_{η}^2 a seguir:
 - $\mu = -5,4$,
 - $\phi \in \{0,50; 0,80; 0,90; 0,99\}$,
 - $\sigma_{\eta}^2 \in \{0,1^2; 0,3^2; 0,5^2\}$.



Aplicação da Metodologia Proposta

- A metodologia proposta foi aplicada num conjunto de dados simulados.
- As séries com 1.461 observações se caracterizam pelo produto cruzado dos possíveis valores predeterminados para μ , ϕ e σ_{η}^2 a seguir:
 - $\mu = -5,4$,
 - $\phi \in \{0,50; 0,80; 0,90; 0,99\}$,
 - $\sigma_{\eta}^2 \in \{0,1^2; 0,3^2; 0,5^2\}$.



Aplicação da Metodologia Proposta

- A metodologia proposta foi aplicada num conjunto de dados simulados.
- As séries com 1.461 observações se caracterizam pelo produto cruzado dos possíveis valores predeterminados para μ , ϕ e σ_{η}^2 a seguir:
 - $\mu = -5,4$,
 - $\phi \in \{0,50; 0,80; 0,90; 0,99\}$,
 - $\sigma_{\eta}^2 \in \{0,1^2; 0,3^2; 0,5^2\}$.



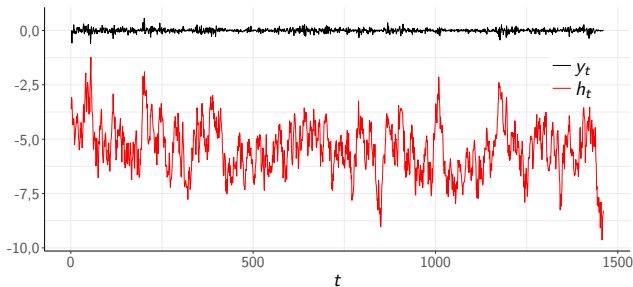


Figura: Conjunto de dados simulados a partir do modelo canônico em (12), $\mu = -5,4$, $\phi = 0,90$ e $\sigma_\eta^2 = 0,5^2$.



Estimação dos Parâmetros

- O algoritmo de estimação foi implementado em C.
- Em todas as execuções foram tomadas:
 - 15.000 iterações;
 - *burn-in* = 5.000;
 - *thin* = 10.
- As distribuições *a priori* de μ , ϕ e σ_η^2 foram definidas como sugere Kim et al. (1998) e Kastner e Frühwirth-Schnatter (2014):

$$\begin{aligned}\mu &\sim \mathcal{N}(0, 100), \\ \phi_0 &\sim \mathcal{B}\left(20, \frac{3}{2}\right), \\ \sigma_\eta^2 &\sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).\end{aligned}\tag{48}$$



Estimação dos Parâmetros

- O algoritmo de estimação foi implementado em C.
- Em todas as execuções foram tomadas:
 - 15.000 iterações;
 - *burn-in* = 5.000;
 - *thin* = 10.
- As distribuições *a priori* de μ , ϕ e σ_η^2 foram definidas como sugere Kim et al. (1998) e Kastner e Frühwirth-Schnatter (2014):

$$\begin{aligned}\mu &\sim \mathcal{N}(0, 100), \\ \phi_0 &\sim \mathcal{B}\left(20, \frac{3}{2}\right), \\ \sigma_\eta^2 &\sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).\end{aligned}\tag{48}$$



Estimação dos Parâmetros

- O algoritmo de estimação foi implementado em C.
- Em todas as execuções foram tomadas:
 - 15.000 iterações;
 - *burn-in* = 5.000;
 - *thin* = 10.
- As distribuições *a priori* de μ , ϕ e σ_η^2 foram definidas como sugere Kim et al. (1998) e Kastner e Frühwirth-Schnatter (2014):

$$\begin{aligned}\mu &\sim \mathcal{N}(0, 100), \\ \phi_0 &\sim \mathcal{B}\left(20, \frac{3}{2}\right), \\ \sigma_\eta^2 &\sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).\end{aligned}\tag{48}$$



Estimação dos Parâmetros

- O algoritmo de estimação foi implementado em C.
- Em todas as execuções foram tomadas:
 - 15.000 iterações;
 - *burn-in* = 5.000;
 - *thin* = 10.
- As distribuições *a priori* de μ , ϕ e σ_η^2 foram definidas como sugere Kim et al. (1998) e Kastner e Frühwirth-Schnatter (2014):

$$\begin{aligned}\mu &\sim \mathcal{N}(0, 100), \\ \phi_0 &\sim \mathcal{B}\left(20, \frac{3}{2}\right), \\ \sigma_\eta^2 &\sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).\end{aligned}\tag{48}$$



Estimação dos Parâmetros

- O algoritmo de estimação foi implementado em C.
- Em todas as execuções foram tomadas:
 - 15.000 iterações;
 - *burn-in* = 5.000;
 - *thin* = 10.
- As distribuições *a priori* de μ , ϕ e σ_η^2 foram definidas como sugere Kim et al. (1998) e Kastner e Frühwirth-Schnatter (2014):

$$\begin{aligned}\mu &\sim \mathcal{N}(0, 100), \\ \phi_0 &\sim \mathcal{B}\left(20, \frac{3}{2}\right), \\ \sigma_\eta^2 &\sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).\end{aligned}\tag{48}$$



Estimação dos Parâmetros

- O algoritmo de estimação foi implementado em C.
- Em todas as execuções foram tomadas:
 - 15.000 iterações;
 - $burn-in = 5.000$;
 - $thin = 10$.
- As distribuições *a priori* de μ , ϕ e σ_η^2 foram definidas como sugere Kim et al. (1998) e Kastner e Frühwirth-Schnatter (2014):

$$\begin{aligned}\mu &\sim \mathcal{N}(0, 100), \\ \phi_0 &\sim \mathcal{B}\left(20, \frac{3}{2}\right), \\ \sigma_\eta^2 &\sim \mathcal{G}\left(\frac{1}{2}, \frac{1}{2}\right).\end{aligned}\tag{48}$$



- Primeiro passo foi avaliar a performance do processo de estimação da variável latente. Para isso, os valores reais de h_t foram comparados através de gráficos com os valores médios dos 1.000 valores estimados de \hat{h}_t :

$$\bar{\hat{h}}_t = \frac{\sum_{j=1}^{1000} \hat{h}_{t,j}}{1000}. \quad (49)$$

- Os possíveis valores do fator de desconto, λ_t , foram limitados ao intervalo (0,75; 1).



- Primeiro passo foi avaliar a performance do processo de estimação da variável latente. Para isso, os valores reais de h_t foram comparados através de gráficos com os valores médios dos 1.000 valores estimados de \hat{h}_t :

$$\bar{\hat{h}}_t = \frac{\sum_{j=1}^{1000} \hat{h}_{t,j}}{1000}. \quad (49)$$

- Os possíveis valores do fator de desconto, λ_t , foram limitados ao intervalo (0,75; 1).



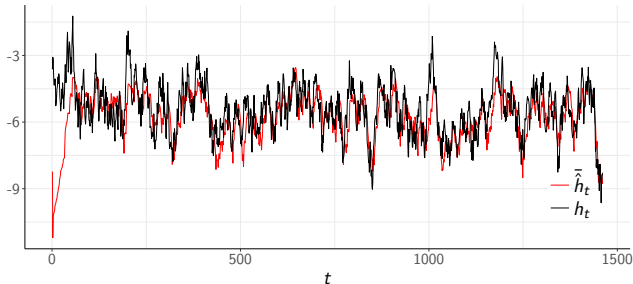
μ, ϕ e σ_η^2 conhecidos

Figura: Valores médios estimados de h_t do modelo (12) quando os parâmetros μ , ϕ e σ_η^2 são conhecidos ($\mu = -5,4$, $\phi = 0,90$ e $\sigma_\eta^2 = 0,5^2$).



Apenas ϕ é conhecido

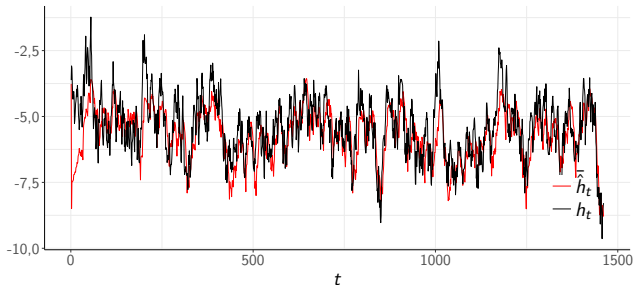


Figura: Valores médios estimados de h_t do modelo (12) quando apenas o parâmetro ϕ é conhecido ($\mu = -5,4$, $\phi = 0,90$ e $\sigma_\eta^2 = 0,5^2$).



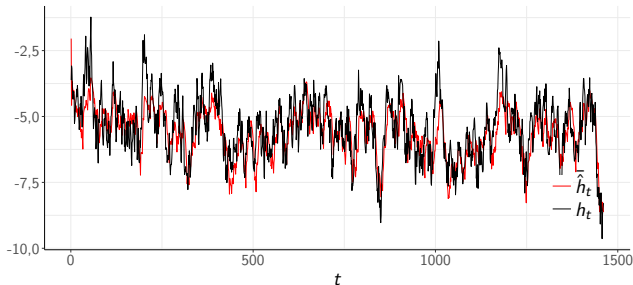
ϕ parte do valor real

Figura: Valores médios estimados de h_t do modelo (12) quando o valor inicial de ϕ é igual ao valor verdadeiro do parâmetro ($\mu = -5,4$, $\phi = 0,90$ e $\sigma_\eta^2 = 0,5^2$).



$$\lambda_t = 0,50$$

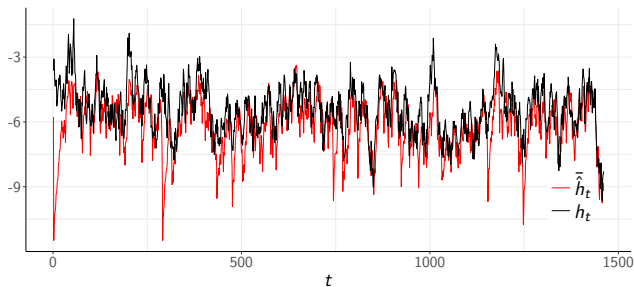


Figura: Valores médios estimados de h_t do modelo (12) quando $\lambda_t = 0,50$ ($\mu = -5,4$, $\phi = 0,90$ e $\sigma_\eta^2 = 0,5^2$).



- O algoritmo foi executado para os 12 conjuntos de dados simulados.
- O valor de λ_t foi fixado próximo ao máximo valor possível em cada caso.
- Os valores de μ são bem estimados em todas as circunstâncias.
- Os valores estimados de σ_η^2 são aceitáveis quando o valor real não é muito pequeno.
- Os valores estimados de ϕ parecem presos num “poço de potencial”.



- O algoritmo foi executado para os 12 conjuntos de dados simulados.
- O valor de λ_t foi fixado próximo ao máximo valor possível em cada caso.
- Os valores de μ são bem estimados em todas as circunstâncias.
- Os valores estimados de σ_η^2 são aceitáveis quando o valor real não é muito pequeno.
- Os valores estimados de ϕ parecem presos num “poço de potencial”.



- O algoritmo foi executado para os 12 conjuntos de dados simulados.
- O valor de λ_t foi fixado próximo ao máximo valor possível em cada caso.
- Os valores de μ são bem estimados em todas as circunstâncias.
- Os valores estimados de σ_η^2 são aceitáveis quando o valor real não é muito pequeno.
- Os valores estimados de ϕ parecem presos num “poço de potencial”.



- O algoritmo foi executado para os 12 conjuntos de dados simulados.
- O valor de λ_t foi fixado próximo ao máximo valor possível em cada caso.
- Os valores de μ são bem estimados em todas as circunstâncias.
- Os valores estimados de σ_η^2 são aceitáveis quando o valor real não é muito pequeno.
- Os valores estimados de ϕ parecem presos num “poço de potencial”.



- O algoritmo foi executado para os 12 conjuntos de dados simulados.
- O valor de λ_t foi fixado próximo ao máximo valor possível em cada caso.
- Os valores de μ são bem estimados em todas as circunstâncias.
- Os valores estimados de σ_η^2 são aceitáveis quando o valor real não é muito pequeno.
- Os valores estimados de ϕ parecem presos num “poço de potencial”.



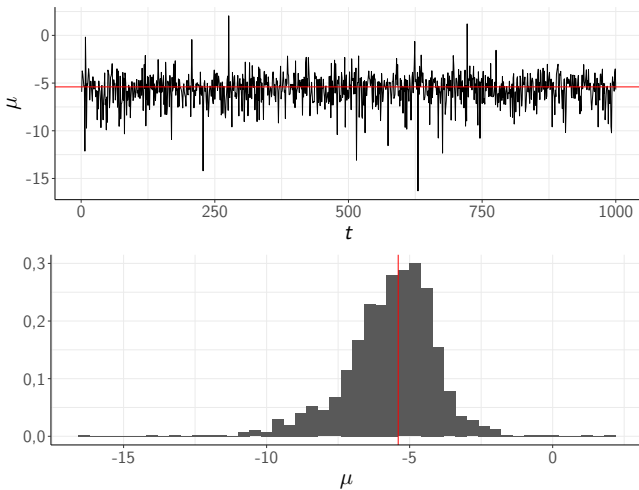


Figura: $\mu = -5,4$, $\phi = 0,99$ e $\sigma_{\eta}^2 = 0,5^2$.



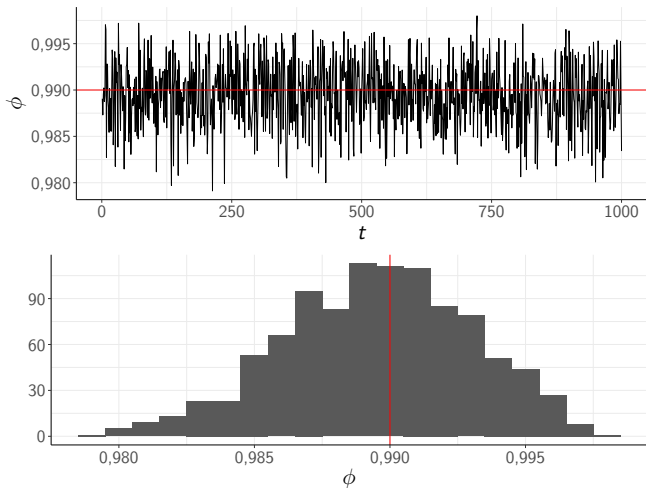


Figura: $\mu = -5,4$, $\phi = 0,99$ e $\sigma_{\eta}^2 = 0,5^2$.



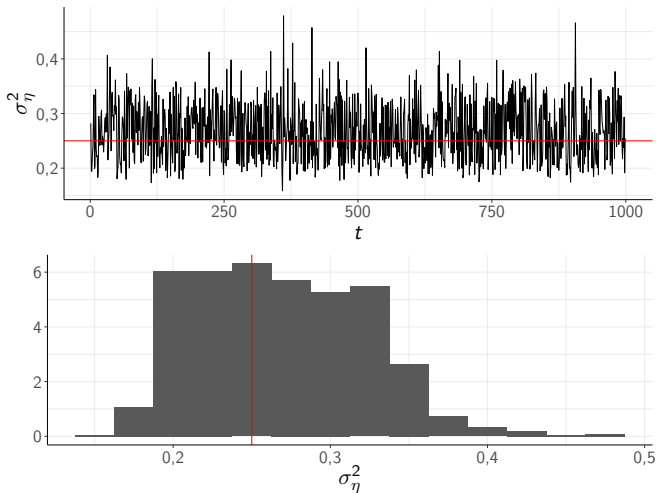


Figura: $\mu = -5,4$, $\phi = 0,99$ e $\sigma_\eta^2 = 0,5^2$.



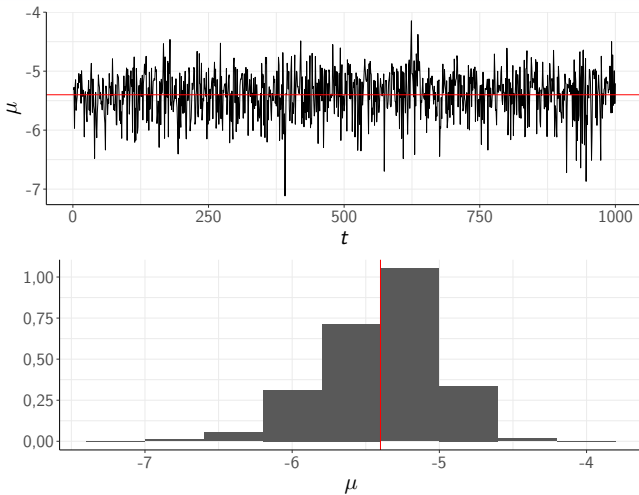


Figura: $\mu = -5,4$, $\phi = 0,80$ e $\sigma_{\eta}^2 = 0,32^2$.



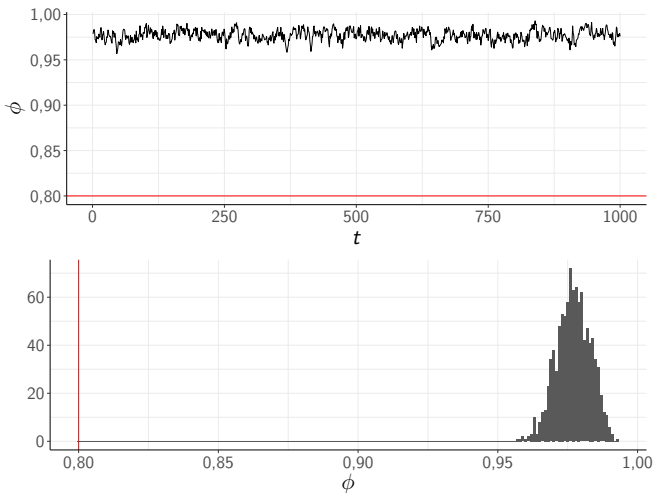


Figura: $\mu = -5,4$, $\phi = 0,80$ e $\sigma_{\eta}^2 = 0,3^2$.



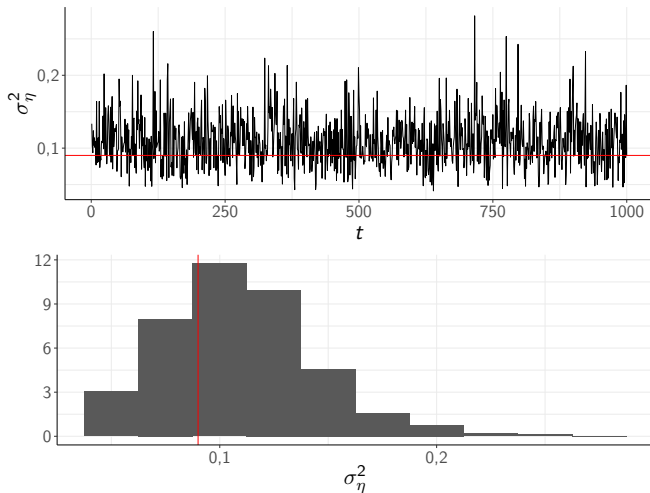


Figura: $\mu = -5,4$, $\phi = 0,80$ e $\sigma_\eta^2 = 0,3^2$.



Comparação dos Resultados com o JAGS

- O modelo de volatilidade estocástica foi implementado em JAGS para comparar com os resultados obtidos a partir da metodologia proposta.
- A escolha do JAGS foi motivada por dois fatores: a) desenvolvimento e manutenção do programa; b) ótima integração com o R.
- Apesar das facilidades computacionais que o JAGS proporciona, os problemas de estimação dos parâmetros persistem e são bem similares.



Comparação dos Resultados com o JAGS

- O modelo de volatilidade estocástica foi implementado em JAGS para comparar com os resultados obtidos a partir da metodologia proposta.
- A escolha do JAGS foi motivada por dois fatores: a) desenvolvimento e manutenção do programa; b) ótima integração com o R.
- Apesar das facilidades computacionais que o JAGS proporciona, os problemas de estimação dos parâmetros persistem e são bem similares.



Comparação dos Resultados com o JAGS

- O modelo de volatilidade estocástica foi implementado em JAGS para comparar com os resultados obtidos a partir da metodologia proposta.
- A escolha do JAGS foi motivada por dois fatores: a) desenvolvimento e manutenção do programa; b) ótima integração com o R.
- Apesar das facilidades computacionais que o JAGS proporciona, os problemas de estimação dos parâmetros persistem e são bem similares.



- De forma semelhante, os valores de μ geraram boas estimativas em todos os casos.
- Os valores estimados de σ_{η}^2 , do mesmo modo, nem sempre foram como esperado. A medida que o valor real de ϕ decresce, o JAGS tende a subestimar os valores da volatilidade.
- Os valores estimados de ϕ não ficaram “presos”, como no caso anterior. Porém as estimativas possuem uma dispersão muito alta, e são bastante dependentes da distribuição *a priori* da persistência.



- De forma semelhante, os valores de μ geraram boas estimativas em todos os casos.
- Os valores estimados de σ_{η}^2 , do mesmo modo, nem sempre foram como esperado. A medida que o valor real de ϕ decresce, o JAGS tende a subestimar os valores da volatilidade.
- Os valores estimados de ϕ não ficaram “presos”, como no caso anterior. Porém as estimativas possuem uma dispersão muito alta, e são bastante dependentes da distribuição *a priori* da persistência.



- De forma semelhante, os valores de μ geraram boas estimativas em todos os casos.
- Os valores estimados de σ_{η}^2 , do mesmo modo, nem sempre foram como esperado. A medida que o valor real de ϕ decresce, o JAGS tende a subestimar os valores da volatilidade.
- Os valores estimados de ϕ não ficaram “presos”, como no caso anterior. Porém as estimativas possuem uma dispersão muito alta, e são bastante dependentes da distribuição *a priori* da persistência.



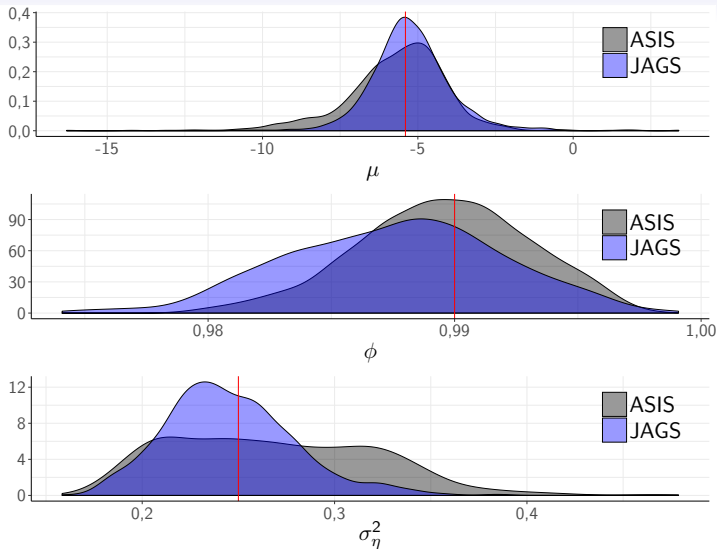


Figura: $\mu = -5,4$, $\phi = 0,99$ e $\sigma^2_\eta = 0,5^2$.



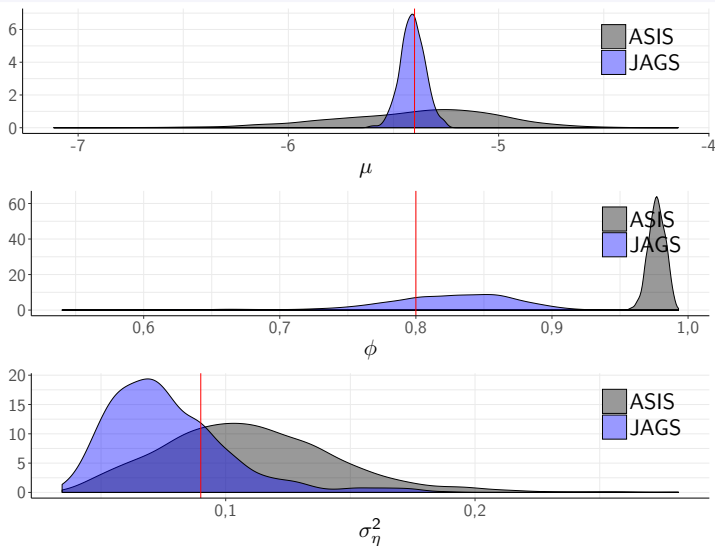


Figura: $\mu = -5,4$, $\phi = 0,80$ e $\sigma^2_\eta = 0,32$.



Um Exemplo com Dados Reais

- Achcar et al. (2011) propõe dois MVE para estudar a concentração de ozônio na Cidade do México.
- Os modelos propostos são multivariados (cinco regiões distintas da cidade).
- Os dados se referem a medições da média semanal do nível diário máximo de ozônio nas cinco regiões.
- O algoritmo proposto, bem como a implementação em JAGS, foram executados para uma região específica (CE).



Um Exemplo com Dados Reais

- Achcar et al. (2011) propõe dois MVE para estudar a concentração de ozônio na Cidade do México.
- Os modelos propostos são multivariados (cinco regiões distintas da cidade).
- Os dados se referem a medições da média semanal do nível diário máximo de ozônio nas cinco regiões.
- O algoritmo proposto, bem como a implementação em JAGS, foram executados para uma região específica (CE).



Um Exemplo com Dados Reais

- Achcar et al. (2011) propõe dois MVE para estudar a concentração de ozônio na Cidade do México.
- Os modelos propostos são multivariados (cinco regiões distintas da cidade).
- Os dados se referem a medições da média semanal do nível diário máximo de ozônio nas cinco regiões.
- O algoritmo proposto, bem como a implementação em JAGS, foram executados para uma região específica (CE).



Um Exemplo com Dados Reais

- Achcar et al. (2011) propõe dois MVE para estudar a concentração de ozônio na Cidade do México.
- Os modelos propostos são multivariados (cinco regiões distintas da cidade).
- Os dados se referem a medições da média semanal do nível diário máximo de ozônio nas cinco regiões.
- O algoritmo proposto, bem como a implementação em JAGS, foram executados para uma região específica (CE).



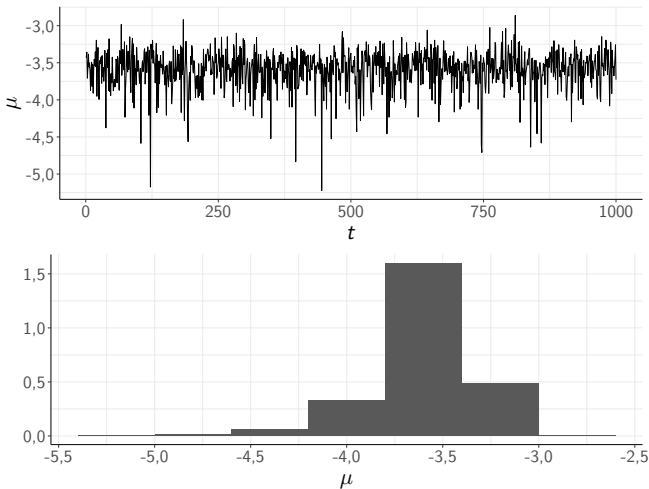


Figura: Valores estimados do MVE aplicado aos dados reais.



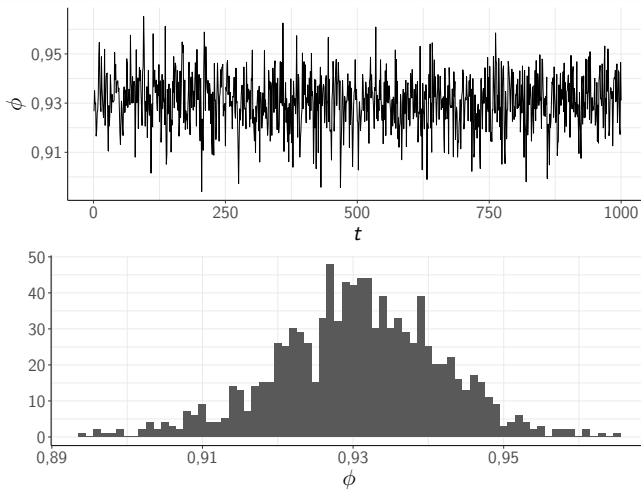


Figura: Valores estimados do MVE aplicado aos dados reais.



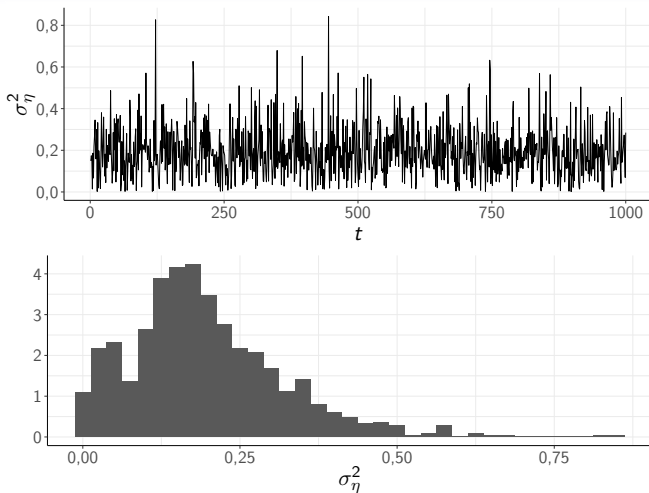


Figura: Valores estimados do MVE aplicado aos dados reais.



Tabela: Estatísticas dos valores estimados dos parâmetros do MVE aplicado aos dados reais.

Parâmetro	Média	2,5%	Mediana	97,5%
μ	-3,594	-4,232	-3,557	-3,197
ϕ	0,931	0,908	0,931	0,951
σ_{η}^2	0,194	0,011	0,178	0,478



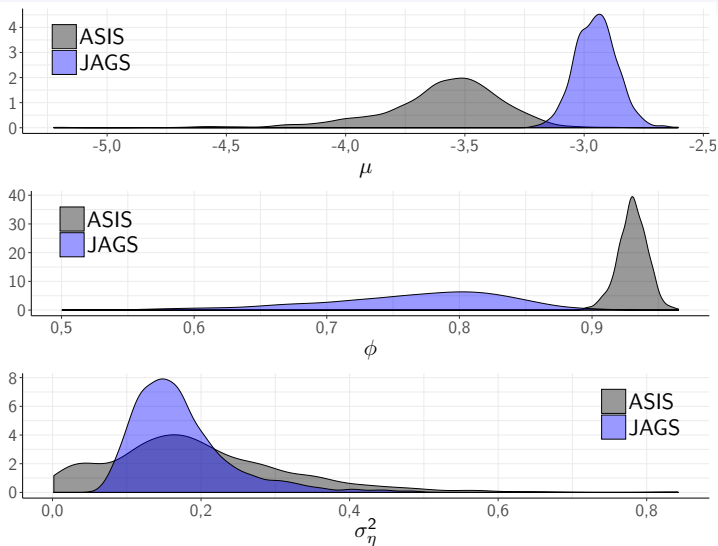


Figura: Valores estimados via JAGS vs método proposto aplicados aos dados reais.



Introdução
○

MLD
○○○○○○

MVE
○○○○○
○○○○○

Estimação
○○○○○○○○
○○○○○○○○

Aplicação
○○○○○○○○○○○○○○
○○○○○○○○○○○○○○

Conclusão
●

Referências



Referências Bibliográficas I

- Achcar, J. A., Rodrigues, E. R., e Tzintzun, G. (2011). Using stochastic volatility models to analyse weekly ozone averages in Mexico City. *Environmental and Ecological Statistics*, 18(2):271–290.
- Bos, C. S. (2012). *Handbook of Volatility Models and their Applications*, chapter Relating Stochastic Volatility Estimation Methods, pages 147–174. John Wiley & Sons.
- Carter, C. K. e Kohn, R. (1994). On Gibbs sampling for state space models. *Biometrika*, 81:541–553.
- Durbin, J. e Koopman, S. J. (1997). Monte Carlo maximum likelihood estimation for non-gaussian state-space models. *Biometrika*, 89:603–615.
- Durbin, J. e Koopman, S. J. (2012). *Time Series Analysis by State Space Methods*. Oxford University Press, Oxford, 2nd edition.



Referências Bibliográficas II

- Gallant, A. R. e Tauchen, G. (1996). Which moments to match? *Econometric Theory*, 12:657–681.
- Harvey, A., Ruiz, E., e Shephard, N. (1994). Multivariate stochastic variance models. *Review of Economic Studies*, 61:247–264.
- Hastings, W. K. (1970). Monte carlo sampling methods using markov chains and their applications. *Biometrika*, 57(1):97–109.
- Kastner, G. e Frühwirth-Schnatter, S. (2014). Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics and Data Analysis*, 76:408–423.
- Kim, S., Shepard, N., e Chib, S. (1998). Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies*, 65(3):361–393.



Referências Bibliográficas III

- McCormick, T. H., Raftery, A. E., Madigan, D., e Burd, R. S. (2012). Dynamic logistic regression and dynamic model averaging for binary classification. *Biometrics*, 68(1):23–30.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. H., e Teller, E. (1953). Equation of state calculations by fast computing machines. *Journal of Chemical Physics*, 21(6):1087–1092.
- Nguyen, T. M. (2007). A new efficient algorithm for the analysis with non-linear and non-gaussian state space models. Master's thesis, VU University Amsterdam.
- Omori, Y., Chib, S., Shephard, N., e Nakajima, J. (2007). Stochastic volatility with leverage: Fast and efficient likelihood inference. *Journal of Econometrics*, 140:425–449.
- Richard, J. F. e Zhang, W. (2007). Efficient high-dimensional importance sampling. *Journal of Econometrics*, 141:1385–1411.



Referências Bibliográficas IV

- Shephard, N. e Pitt, M. K. (1997). Likelihood analysis of non-gaussian measurement time series. *Biometrika*, 84:653–667.
- Taylor, S. (1982). Financial returns modelled by the product of two stochastic processes – a study of daily sugar prices 1961-75. *Time Series Analysis: Theory and Practice*, 1:203–226.
- Tierney, L. e Kadane, J. B. (1986). Accurate approximations for posterior moments and marginal densities. *Journal of the American Statistical Association*, 81(393):82–86.
- West, M. e Harrison, J. (1997). *Bayesian Forecasting and Dynamic Models*. Springer, New York, 2nd edition.
- Yu, Y. e Meng, X. (2011). To center or not to center: That is not the question - an ancillarity-sufficiency interweaving strategy (asis) for boosting mcmc efficiency. *Journal of Computational and Graphical Statistics*, 20(3):531–570.



Referências Bibliográficas V

Zivot, E. e Yollin, G. (2012). Time series forecasting with state space models. In *R/Finance Workshop, University of Chicago*.

