econometrics

MDPI

Article

Multivariate Stochastic Volatility Modeling via Integrated Nested Laplace Approximations: A Multifactor Extension

João Pedro Coli de Souza Monteneri Nacinben † and Márcio Laurini *,†

Department of Economics, FEARP-University of São Paulo, Ribeirão Preto 14040-905, Brazil; joao.nacinben@alumni.usp.br

* Correspondence: laurini@fearp.usp.br; Tel.: +55-16-3329-0867

† These authors contributed equally to this work.

Abstract: This study introduces a multivariate extension to the class of stochastic volatility models, employing integrated nested Laplace approximations (INLA) for estimation. Bayesian methods for estimating stochastic volatility models through Markov Chain Monte Carlo (MCMC) can become computationally burdensome or inefficient as the dataset size and problem complexity increase. Furthermore, issues related to chain convergence can also arise. In light of these challenges, this research aims to establish a computationally efficient approach for estimating multivariate stochastic volatility models. We propose a multifactor formulation estimated using the INLA methodology, enabling an approach that leverages sparse linear algebra and parallelization techniques. To evaluate the effectiveness of our proposed model, we conduct in-sample and out-of-sample empirical analyses of stock market index return series. Furthermore, we provide a comparative analysis with models estimated using MCMC, demonstrating the computational efficiency and goodness of fit improvements achieved with our approach.

Keywords: multivariate stochastic volatility; integrated nested laplace approximations; Bayesian methods; computational efficiency

check for updates

Citation: Nacinben, João Pedro Coli de Souza Monteneri, and Márcio Laurini. 2024. Multivariate Stochastic Volatility Modeling via Integrated Nested Laplace Approximations: A Multifactor Extension. *Econometrics* 12: 5. https://doi.org/10.3390/econometrics12010005

Academic Editor: Olivier Parent

Received: 5 December 2023
Revised: 14 February 2024
Accepted: 17 February 2024
Published: 19 February 2024

1. Introduction

In economics, especially in the field of finance, time series modeling recurrently involves incorporating patterns of dependence in the conditional variance. In financial series, the so-called conditional volatility models are of particular interest, as variance plays a considerable role in determining asset prices, measuring risk and in the construction of tools to hedge them. By understanding the underlying factors driving the volatility dynamics, policymakers can implement measures to mitigate systemic risks and enhance financial stability. This may involve implementing macroprudential policies aimed at reducing excessive volatility and preventing financial crises.

However, as it is a latent process, conditional volatility is not easily estimated using conventional tools, which would include, for example, maximum likelihood estimators. In response, the literature presents the development of different paths regarding the treatment given to the assumed volatility process.

Among the proposed methods, one notable class is the family of univariate ARCH models (Autoregressive Conditional Heteroskedasticity), originally proposed by Engle (1982) and generalized into the univariate GARCH class (Generalized Autoregressive Conditional Heteroskedasticity) by Bollerslev (1986).

Another significant class of models treats conditional volatility as a stochastic process, commonly referred to as stochastic volatility (SV) models. Introduced by Taylor (1986) using a nonlinear state-space representation with the log-variance following an AR(1) process, univariate SV models offer certain advantages over the univariate ARCH class. They do not necessitate the assumption of a deterministic structure for latent variance,

and using autoregressive formulations for the conditional variance, they allow for a more straightforward multivariate extension using formulations based on vector autoregressive (VAR) models and common factor structures. However, SV models come with greater complexity in their estimation due to the presence of latent variables in the likelihood of the process.

In the frequentist perspective, significant development centers around the quasi-maximum likelihood method based on prediction error decomposition via the Kalman Filter, independently proposed by Nelson (1988) and Harvey et al. (1994). Alternatively, Bayesian estimation techniques are particularly interesting because latent processes can be treated as additional parameters to be estimated. Refer to Kim and Shephard (1998) for the original Bayesian inference of SV models and, more recently, Kastner and Frühwirth-Schnatter (2014), both of which rely on Markov Chain Monte Carlo (MCMC) algorithms.

From an applied standpoint, simulation-based methodologies like MCMC can pose computational challenges, especially as the volume of data involved in the estimation grows in terms of dimensionality and the number of observations. These algorithms are susceptible to chain convergence issues, as highlighted by Rajaratnam and Sparks (2015), in addition to facing computational inefficiency due to the iterative and non-parallelizable nature of the method, resulting in longer execution times. Consequently, stochastic volatility models may become less appealing or even computationally prohibitive when dealing with very long series or multivariate formulations.

Martino et al. (2011) propose an alternative approach to Bayesian estimation of stochastic volatility models using integrated nested Laplace approximations (INLA), originally introduced by Rue et al. (2009). This approach allows for the estimation of parameters and latent variables through precise deterministic approximations to the posterior distributions, provided that the model can be approximated by a Gaussian Markov random field. Notably, Chaim and Laurini (2019) highlights that the INLA method, based on an analytical approach, eliminates the need for simulation procedures, rendering it immune to issues related to chain convergence.

As the formulation used can be represented as a Gaussian Markov Random Field, it allows a representation using a conditional Markov structure (Rue and Held 2005). The Markov property allows for representing the model's precision matrix with a sparse matrix, which allows for the use of sparse linear algebra with relevant computational gains. Another important point is that Laplace approximations are numerically parallelizable, which allows all the computational power to be exploited using multiple processors.

The literature using this methodology has primarily focused on the analysis of the univariate case, where a single asset is considered in the state space formulation. However, extending the estimation of multivariate stochastic volatility models via integrated nested Laplace approximations is a natural progression, considering that works like Martino et al. (2011) only examine the bivariate case. Nevertheless, the dimensionality of the problem poses critical implications for the use of INLA in these models. As noted by Martino et al. (2011), the informational gain (greater volume of data and additional dimension) may lead to a substantial increase in estimation time, potentially undermining the efficiency of this tool.

One possible solution is to search for a more parsimonious parameterization, which allows the number of parameters to be estimated not to grow too much as the number of series analyzed also increases. Factor models fulfill exactly this role, as Asai et al. (2006) observes.

This study aims to develop a practical framework for estimating multivariate stochastic volatility models using a shared factors structure. The proposed formulation is based on a structure of common factors, where the log-volatility of each series is given by the combination of latent factors through a structure of estimated factor loadings. This structure can be implemented directly using the INLA methodology, taking advantage of all the computational gains related to the use of integrated Laplace approximations for parameters and latent factors, and also through the representation of sparse matrices and parallelism in the evaluation of the model that is possible through the Gaussian Markov Random Field

representation of the proposed model. Through in-sample and out-of-sample analyses, we show the computational, fitting and forecasting gains derived from the proposed specification, compared to univariate specifications and the alternative multifactor formulation proposed by Kastner et al. (2017).

Our empirical focus centers on analyzing daily returns from stock exchange indices across various global regions. The primary objective is to uncover volatility interdependencies within these specified datasets while evaluating potential accuracy improvements compared to conventional univariate models. Additionally, we seek to assess the computational efficiency of our approach in comparison to commonly used MCMC algorithms.

The main contribution of the method is an efficient computational implementation for factor models of stochastic volatility, bypassing the computational cost and possible convergence problems existing in Bayesian estimations using MCMC. In this way, the proposed method can be used as an efficient computational tool for estimating risk, measured by conditional variance, in multivariate problems, with applications in portfolio allocation, construction of tail risk measures and hedging procedures.

This work is organized into five sections, beginning with this introductory section. Section 2 provides a comprehensive review and discussion of existing literature concerning the estimation of stochastic volatility models, and the use of INLA (Integrated Nested Laplace Approximations), with a particular emphasis on the multivariate aspect within the context of factor modeling. In Section 3, we offer a detailed exposition of the methodology to be employed, accompanied by an introduction to the datasets that will underpin the evaluation of our proposed models. The results of our analysis are presented in Section 4, with in- and out-of-sample analysis using real data, preceding the final Section 5 where we succinctly summarize the key findings.

2. Literature Review

Our exploration begins with the early developments in the estimation of SV models, followed by an examination of the integration of Bayesian methodologies and the application of the integrated nested Laplace approximation (INLA) method. Concluding this section, we provide an overview of the literature pertaining to multivariate and factorial models.

2.1. Stochastic Volatility Models

In general, the literature on stochastic volatility, especially concerning its theoretical foundations and principles, is well-established. Recent advancements are primarily associated with estimation methods. Taylor (1986) introduced the concept of stochastic volatility in the univariate case, assuming log-normal returns and a latent process of (log-)volatility. The unobservable nature of this process initially led to a search for methodological developments to overcome the limitations of traditional techniques, such as maximum likelihood. In this context, Bayesian inference techniques, in conjunction with Markov Chain Monte Carlo (MCMC) methods, gained popularity, fostering dedicated literature aimed at proposing more efficient alternatives. One such alternative is integrated nested Laplace approximations (INLA), introduced by Rue et al. (2009) and applied to stochastic volatility by Martino et al. (2011). This framework appears to offer accuracy at least on par with MCMC, as evidenced by Ehlers and Zevallos (2015), while also providing advantages in terms of estimation efficiency.

Starting from a sequence $\{r_t\}_{t=1}^{T}$ of returns, the SV model of Taylor (1986) model can be represented as follows:

$$r_t = \exp\{h_t/2\}\varepsilon_t, \quad \varepsilon_t \sim N(0,1) \tag{1}$$

$$h_t = \mu + \phi(h_{t-1} - \mu) + \eta_t, \quad \eta_t \sim N(0, [(1-\phi^2)\tau_h]^{-1}) \tag{2}$$

Log-variance, h_t, is assumed to be a stationary first-order autoregressive process, whose dynamics are configured by the parameters μ (long-term average), ϕ (related to autoregressive persistence) and τ_h (marginal precision of the log-variance of h_t).

An immediate challenge arises from the unobservable nature of log-volatility. Following Taylor's (1986) formulation, it is suggested to estimate μ, ϕ, and τ $(= \tau_h[1-\phi^2])$ using sample moments, but not h_t. The impossibility of employing a maximum likelihood estimator is noted since it would necessitate dealing with a multiple integral with a dimension equal to the sample size to marginalize the latent variance vector. Harvey (1989) proposed a frequentist approach to address the issue of estimating the latent variable in the model. This approach involves a quasi-maximum likelihood estimator (referred to as "quasi-maximum likelihood") based on prediction error decomposition via the Kalman Filter. To achieve this, the square of the log-returns is linearized, involving an average decomposition along with a first-order autoregressive process, thereby imposing a linear representation of the state space for the model.

Alternatively, Sandmann and Koopman (1998) discussed a methodology that replaces quasi-likelihood with a simulated version, called Monte Carlo Maximum Likelihood (MCML), which also allows for the application of the Kalman Filter in the decomposition of the prediction error. Andersen et al. (1999), on the other hand, returned to the class of estimators that aim to explore the moments of the sample, using the Efficient Method of Moments (EMM), in which the derivative of the log-likelihood function (score vector) provides the moment conditions. In the authors' study, the EMM estimator achieves the efficiency of maximum likelihood estimators in large samples.

2.2. Multivariate and Factor Models

Since the inception of stochastic volatility models first proposed by Taylor (1986), several extensions have emerged, introducing a wider array of specifications and facilitating the multivariate analysis of conditional volatility patterns. Building upon the work of Harvey et al. (1994), multivariate SV modeling has gained momentum and undergone further developments, incorporating appropriate functional forms to capture specific stylized facts and presenting a more efficient approach to estimation and parameterization.

Amid these developments, the concept of factor structures has been introduced into the realm of multivariate stochastic volatility models, with greater emphasis given by Jacquier et al. (1995). These factor structures addressed the computational complexity associated with the high dimensionality inherent to multivariate volatility modeling.

Harvey et al. (1994) laid the groundwork for what they referred to as a multivariate generalization of stochastic variance models. Drawing from the multivariate representation of ARCH models and the associated restrictions proposed by Bollerslev (1990), Harvey et al. (1994) assumed a set of equations to describe the model:

$$r_{it} = \varepsilon_{it}(\exp\{h_{it}\})^{1/2}, \quad i = 1, \ldots, N; \quad t = 1, \ldots, T \tag{3}$$

$$h_{it} = \gamma_i + \varphi_i h_{it-1} + z_{it} \tag{4}$$

where r_{it} is the observation of the i-th series in the period t and the vectors $\varepsilon_t = (\varepsilon_{1t}, \ldots, \varepsilon_{Nt})'$ and $z_t = (z_{1t}, \ldots, z_{Nt})'$ of errors follow a multivariate normal distribution with zero mean and covariance matrices Σ_ε and Σ_η, respectively. It is also assumed that Σ_ε has each element of its main diagonal equal to 1 and its other entries denoted by ρ_{ij}. For estimation, the authors adopt a frequentist strategy, using a quasi-maximum likelihood estimator by means of the Kalman filter.

From a Bayesian perspective, the work carried out by Martino (2007) and Martino et al. (2011) introduced an estimation scheme based on the INLA methodology for the multivariate case. As opposed to the univariate formulation, adopted in the vast majority of works in which INLA is used, the multivariate model is capable of capturing a greater variety of data characteristics, for example, the spillover volatility effect, in which knowledge about

one asset could help make predictions about another. Martino (2007) proposed a bivariate extension, written as

$$\mathbf{r_t} = \mathbf{\Omega_t}\boldsymbol{\varepsilon_t}, \quad \boldsymbol{\varepsilon_t} \sim N(\mathbf{0}, \mathbf{\Sigma}_\varepsilon) \tag{5}$$

$$\mathbf{h_t} = \boldsymbol{\nu} + \mathbf{\Phi}(\mathbf{h_{t-1}} - \boldsymbol{\nu}) + \boldsymbol{\eta_t}, \quad \boldsymbol{\eta_t} \sim N(\mathbf{0}, \mathbf{\Sigma}_\eta) \tag{6}$$

where $\mathbf{r_t} = \{r_{t1}, r_{t2}\}$ are the observed log-returns on t, $\boldsymbol{\varepsilon_t} = \{\varepsilon_{t1}, \varepsilon_{t2}\}$ and $\boldsymbol{\eta_t} = \{\eta_{t1}, \eta_{t2}\}$ are bivariate noise terms and $\mathbf{h_t} = \{h_{t1}, h_{t2}\}$ are the latent volatilities. Furthermore,

$$\mathbf{\Phi} = \begin{pmatrix} \phi_{11} & \phi_{12} \\ \phi_{21} & \phi_{22} \end{pmatrix} \quad \mathbf{\Omega_t} = \begin{pmatrix} \exp\{h_{1t}/2\} & 0 \\ 0 & \exp\{h_{2t}/2\} \end{pmatrix} \tag{7}$$

$$\mathbf{\Sigma}_\eta = \begin{pmatrix} 1/\tau_{\eta_1} & \rho_\eta/\sqrt{\tau_{\eta_1}\tau_{\eta_2}} \\ \rho_\eta/\sqrt{\tau_{\eta_1}\tau_{\eta_2}} & 1/\tau_{\eta_2} \end{pmatrix} \quad \mathbf{\Sigma}_\varepsilon = \begin{pmatrix} 1 & \rho_\varepsilon \\ \rho_\varepsilon & 1 \end{pmatrix} \tag{8}$$

A generalization to larger dimensions was presented by Shapovalova (2021), who noted that this methodology might lose some of its computational efficiency advantages when applied to higher-order multivariate specifications.

Beyond the basic model, a variety of functional forms and specifications are possible, each making different assumptions about the correlation between volatilities and imposing distinct restrictions and parameterizations. Asai et al. (2006) identified and categorized these variants into four groups: asymmetric models, time-varying correlation models, factor models, and alternative specification models.

The first category of Multivariate Stochastic Volatility (MSV) models incorporates the asymmetries observed in the behavior of asset returns. Specifically, it accounts for the tendency for negative and positive variations to have different impacts on the volatility of the series, a concept known as the leverage effect, first identified by Black (1976). This phenomenon, which deals with a negative correlation between volatility and past returns was introduced to the stochastic volatility context by Harvey and Shephard (1996) and explored within the context of multivariate modeling by Danıelssonn (1998).

A second group of MSV models aimed to capture conditional correlations subject to variations over time, relaxing the assumption of constant correlations assumed in Harvey et al.'s (1994) basic model. One of the ways to do this, as proposed by Yu and Meyer (2006), is to use a Fisher transform for the correlation parameter.

So far, the categories we have discussed have been primarily driven by the goal of enhancing modeling flexibility rather than advancing computational efficiency in model estimation. Factor models, on the other hand, are designed to provide more parsimonious formulations with regard to the number of parameters involved in the problem. Asai et al. (2006) classified this category of Multivariate Stochastic Volatility (MSV) models into two types: multiplicative and additive models.

The first additive factor model can be traced back to Harvey et al. (1994) and was later expanded upon by Jacquier et al. (1995). This model begins with a factor-based structure for the $(m \times m)$ covariance matrix and k active factors. The problem is formulated as follows:

$$\mathbf{r_t} = \mathbf{\Delta} \boldsymbol{F_t} + \Sigma^{-\frac{1}{2}} \epsilon_t \tag{9}$$

$$\mathbf{F_t} \sim N_k(\mathbf{0}, \mathbf{H_t}) \tag{10}$$

$$\log h_{it} = \alpha_i + \delta_i \log h_{i,t-1} + \sigma_{iv} \mathbf{v_t}, \tag{11}$$

In these equations, $\mathbf{H_t} = \text{diag}(\mathbf{h_t})$, with $\mathbf{h_t}$ denoting the k univariate h_{it} processes, and $\mathbf{\Delta} = (\Delta_1, \ldots, \Delta_k)$ represents a $(m \times k)$ matrix of factor loadings. The error term ϵ_t follows a normal distribution with zero mean and an identity covariance matrix of order m, with parameters given by Σ. The authors estimated this model using Bayesian techniques, breaking down the posterior distribution $\pi(\mathbf{F}, \mathbf{\Delta}, \omega|\mathbf{y})$ into conditional distributions and employing a Metropolis–Hastings Markov Chain Monte Carlo (MCMC) sampling algorithm.

When compared to the basic MSV model by Harvey et al. (1994), the additive factor model stands out due to the fact that the number of parameters to be estimated grows only linearly as more return series are introduced into the analysis, i.e., as the dimension increases. In fact, Asai et al. (2006) demonstrated that, for the one-factor case, there are a total of $5m - 1$ parameters, where m represents the number of series. This makes the estimation of multivariate stochastic volatility models using methods involving simulation and chain convergence less computationally expensive.

Another aspect of factor models that moved in the same direction is the class of multiplicative models, partly inspired by the work of Quintana and West (1987). According to Asai et al. (2006), this class of MSV model involves separating log-returns into two components: a noise vector and a common factor.

Asai et al. (2006) also discusses alternative specifications. The authors enumerate four alternative formulations for MSV models which rely on exponential matrix transformation, Cholesky decomposition, Wishart autoregressive process, and the observed range. In the first two cases, the motivation is to guarantee the construction of positive matrices defined for the covariance Σ_t, through the attribution of matrix exponential properties and Cholesky decomposition.

Another important category of approximate Bayesian inference methods is rooted in Variational Bayes (VB) techniques Tan and Nott (2018). VB frames the task of estimating the posterior distribution as an optimization problem, approximating the posterior density with a simpler distribution characterized by unknown parameters. This simplification often takes the form of a multivariate Gaussian distribution with an unknown mean and covariance matrix. Such an approach offers substantial computational advantages over traditional Markov Chain Monte Carlo (MCMC) methods. Moreover, VB can be integrated with sequential methods, facilitating parameter and latent state updates as new data points are observed. The application of Variational Bayes in the estimation of factor Stochastic Volatility (SV) models has been proposed in Gunawan et al. (2021), showcasing notable improvements in computational efficiency along with favorable fitting and prediction performance compared to MCMC-based approaches.

2.3. Bayesian Estimation and Integrated Nested Laplace Approximations (INLA)

In parallel with frequentist developments, several Bayesian estimation techniques have been proposed, capitalizing on the flexibility of Bayesian inference methods when dealing with models that include latent variables treated as additional parameters. However, MCMC approaches can encounter typical challenges such as slow chain convergence, owing to the high correlation among components of the latent volatility variable.

In addition to convergence issues, MCMC can also suffer from computational inefficiencies when applied to very large samples. Martino et al. (2011) suggested an alternative approach for estimating stochastic volatility models using integrated nested Laplace approximations (INLA), a methodology originally introduced by Rue et al. (2009). As long as models can be approximated or represented as Gaussian Markov random fields (GMRF) (see Rue and Held (2005) for the definition and properties of GMRF) both parameters and latent factors can be estimated. The advantage of this approach lies in its utilization of analytical calculations (eliminating the need for simulations) with sparse matrices and operations that are easily parallelizable.

As demonstrated by Rue et al. (2009), through conditional representation, the precision matrix $\mathbf{Q}$ is typically sparse, with only $\mathcal{O}(n)$ of the n^2 entries being non-zero in most cases. Furthermore, it can be factorized as $\mathbf{L}\mathbf{L}^T$, where $\mathbf{L}$ represents the lower Cholesky triangle, inheriting the sparsity of the precision matrix.

Starting with the standard model of stochastic volatility, which includes the equations for r_t and h_t, the proposition of Martino et al. (2011) is extended by assigning a Gaussian prior distribution to the mean parameter μ, with zero mean and known variance. Conse-

quently, the standard model can be interpreted as a latent Gaussian model in which the latent field $\mathbf{x}$ is defined as:

$$\mathbf{x} = \{h_1, ..., h_n, \mu\} \sim N(0, \mathbf{Q}^{-1}(\boldsymbol{\theta_1})), \tag{12}$$

and $\boldsymbol{\theta_1} = \{\tau_h, \phi\}$ are parameters behind volatility. The latent Gaussian field determined by $|\mathbf{x}|$ is partially observed through the conditionally independent data set $\mathbf{r} = \{r_1, ..., r_n\}$, where n denotes the number of observations, with likelihood

$$\pi(\mathbf{r}|\mathbf{x}, \boldsymbol{\theta_2}) = \prod_{t=1}^{n} \pi(r_t|\mathbf{h}, \boldsymbol{\theta_2}), \tag{13}$$

and $\boldsymbol{\theta_2}$ are process parameters $\varepsilon_t \sim N(0, 1)$. Considering $\boldsymbol{\theta} = \{\boldsymbol{\theta_1}, \boldsymbol{\theta_2}\}$, stochastic volatility models can be evaluated by calculating the marginal distributions

$$\pi(\mathbf{x}, \boldsymbol{\theta}|\mathbf{r}) \propto \pi(\boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta}) \prod_{t=1}^{n} \pi(r_t|h_t, \boldsymbol{\theta}). \tag{14}$$

The INLA approach can then be used to perform inference on marginals of $\pi(\mathbf{x}, \boldsymbol{\theta}|\mathbf{r})$, allowing accurate approximations for $\pi(h_t|\mathbf{r})$, $\pi(\mu|\mathbf{r})$ and $\pi(\theta_j|\mathbf{r})$, due to its Gaussian approximation for densities given by the form

$$\pi(\mathbf{x}|\mathbf{r}, \boldsymbol{\theta}) \propto \exp\{-\frac{1}{2}\mathbf{x}^T\mathbf{Q}\mathbf{x} + \sum g_t(h_t)\}, \tag{15}$$

where $g_t(h_t) = \log \pi(r_t|h_t, \boldsymbol{\theta})$. In the context of stochastic volatility models and the formulations discussed by Martino et al. (2011), the precision matrix $\mathbf{Q}$ is tridiagonal, in addition to having the last row and the last column with non-zero values. Still, the vast majority of entries in this matrix continue to consist of zeros, so that $\mathbf{Q}$ is sparse.

Given this structure, the model can be estimated using the INLA methodology proposed in Rue et al. (2009). For space reasons, we do not present the INLA methodology, but details can be obtained in Rue et al. (2009) and recent developments in the implementation of the method in Van Niekerk et al. (2023)

In general, there is a consensus regarding the speed advantages of INLA over MCMC, as explicitly measured in studies like Rue et al. (2009) and Chaim and Laurini (2019). In the former work, taking into account code compilation time, INLA was approximately eight times faster in each estimation compared to MCMC. In the latter study, which analyzed the computational cost of SV models, the computation gains were on the order of 25 times, marking a substantial difference. In terms of estimation accuracy, Ehlers and Zevallos (2015) found improvements associated with the INLA method, particularly when calculating a Value-at-Risk (VaR) metric, comparing it with a quasi-maximum likelihood estimator. Similarly, Chaim and Laurini (2019) investigated the method's accuracy for series with long memory, this time in comparison to traditional MCMC techniques, and found comparable performance between the two.

Furthermore, there has been a growing body of literature focusing on the practical applications of the INLA methodology, with an emphasis on implementations in the R language, using the R-INLA package[1]. Among these contributions, Ruiz-Cárdenas et al. (2012) introduced a general computational approach for Bayesian inference using R-INLA for time series models, expanding the range of dynamic models accessible through the tool. Martins et al. (2013) documented some of the functionalities and resources added to the package since its initial versions. Specifically, within the context of stochastic volatility models, Ravishanker et al. (2022) has provided practical estimation guides and code examples for execution in R, covering modeling with both Gaussian errors and heavy-tailed distributions, such as Student's t distribution.

3. Proposed Model and Methodological Aspects

Seeking to achieve the objectives of establishing a feasible and efficient way of estimating multivariate stochastic volatility models, we adapt the methodology based on integrated nested Laplace approximations for the case of multifactor formulations. In principle, such models can be built using formulations close to the univariate case, increasing the number of observation (state) equations and incorporating latent factor structures.

3.1. Benchmark and Starting Point

The work starts from a specification similar to that presented by Kastner et al. (2017) and Hosszejni and Kastner (2021), which served as a benchmark model. The authors build a structure for efficient estimation via MCMC of a model with m series and k factors. They are defined as

$$\mathbf{r_t}|\boldsymbol{\beta}, \boldsymbol{\Lambda}, \mathbf{f_t}, \bar{\Sigma}_t \sim \mathcal{N}_m(\boldsymbol{\beta} + \boldsymbol{\Lambda}\mathbf{f_t}, \bar{\Sigma}_t) \tag{16}$$

$$\mathbf{f_t}|\tilde{\Sigma}_t \sim \mathcal{N}_r(\mathbf{0}, \tilde{\Sigma}_t) \tag{17}$$

$\boldsymbol{\beta} = (\beta_1, \ldots, \beta_m)'$ is a vector with the mean parameters, $\mathbf{f_t} = (f_{1t}, \ldots, f_{kt})'$ is a vector of latent factors and $\boldsymbol{\Lambda}$ is a $m \times k$ matrix of factor loadings. The following diagonal covariance matrices are also assumed:

$$\bar{\Sigma}_t = \begin{pmatrix} \exp\{\bar{h}_{1t}\} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \exp\{\bar{h}_{mt}\} \end{pmatrix}, \quad \tilde{\Sigma}_t = \begin{pmatrix} \exp\{\tilde{h}_{1t}\} & \ldots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \ldots & \exp\{\tilde{h}_{rt}\} \end{pmatrix}, \tag{18}$$

$$\bar{h}_{it} \sim \mathcal{N}(\bar{\mu}_i + \bar{\phi}_i(\bar{h}_{i,t-1} - \bar{\mu}_i), \bar{\sigma}_i^2), \quad i = 1, \ldots, m, \tag{19}$$

$$\tilde{h}_{jt} \sim \mathcal{N}(\tilde{\mu}_j + \tilde{\phi}_j(\tilde{h}_{j,t-1} - \tilde{\mu}_j), \tilde{\sigma}_j^2), \quad j = 1, \ldots, r. \tag{20}$$

For identification reasons concerning the time-varying covariance matrix[2] (Σ_t), we impose some restrictions on parameters, such as setting the level of log-variances to zero, that is, $\tilde{\mu}_j = 0$ $(j = 1, \ldots, r)$. In addition, Kastner et al. (2017) frees the loading matrix $\boldsymbol{\Lambda}$, in order to identify only sign changes between its elements. See Frühwirth-Schnatter et al. (2023) for further discussion of identification structures in factor models.

To proceed with the estimation, Kastner et al. (2017) defined the prior distributions for the mean parameters, unobserved variances and factor loadings. In the case of $\boldsymbol{\beta}$, Gaussian distributions are chosen, with $\beta_j \sim N(\mathbf{b}_\beta, \mathbf{B}_\beta)$, where $\mathbf{b}_\beta$ is a vector of means and $\mathbf{B}_\beta$ is a variance-covariance matrix. Regarding the latent variance parameters, notably the persistence of the series ($\bar{\phi}_i$ and $\tilde{\phi}_j$), a prior of type $(\phi + 1)/2 \sim \mathcal{B}(a_\phi, b_\phi)$, that is, the Beta distribution. The choice is justified by the objective of avoiding non-stationarity in the variance process, which implies limiting ϕ to the interval $(-1, 1)$. According to Hosszejni and Kastner (2021), since financial series tend to present very persistent variances, with values of ϕ close to 1, it is possible to establish an informative prior distribution by choosing $a_\phi \geq 5$ and $b_\phi \approx 1.5$, such that higher values for ϕ would be more likely.

The distributions for the log-variance volatility, σ, and its initialization parameter, h_0, are also specified. The *gamma* and normal distributions are used, respectively, in such a way that:

$$\sigma^2 \sim \mathcal{G}(\frac{1}{2}, \frac{1}{2B_\sigma}) \tag{21}$$

$$h_0 \sim N(\mu, \frac{\sigma^2}{(1 - \phi^2)}), \tag{22}$$

assuming **h** stationary, an assumption that can be relaxed by defining a prior of the type $h_0 \sim N(\mu, B_h)$, with B_h representing a constant variance. Finally, regarding the prior behind the loading matrix $\boldsymbol{\Lambda}$, independent and zero-centered Gaussian distributions are adopted ($\Lambda_{ij} \sim N[0, B_\Lambda]$), and are an important point for identifying the factor variance, as reported by Kastner et al. (2017).

The estimation of parameters through MCMC sampling employs a Metropolis–Hastings scheme, bolstered by interweaving strategies from *ASIS*[3], aimed at mitigating issues related to slow chain convergence. Practical implementation can be carried out in the *R* programming language using the factorstochvol package (See Hosszejni and Kastner 2021). This package facilitates the easy configuration of hyperparameters, model identification structure, and sampler adjustments, as well as enabling predictions within and outside the observed sample. The sampler is coded in *C++*, offering superior execution speed for the draw stage (*draws*), which is typically computationally intensive.

From a theoretical standpoint, Kastner et al. (2017) presented an algorithm for conditional MCMC sampling in four steps, commencing with the selection of initial values for μ_i, ϕ_i, σ_i, $\mathbf{\Lambda}$, $\mathbf{h}$, and $\mathbf{f}$. In the first stage, the m idiosyncratic variances $\mathbf{h}_{i,\bullet} = (h_{i0}, \ldots, h_{iT})'$ are estimated, still in the univariate context, alongside the parameters μ_i, ϕ_i, and σ_i ($i = 1, \ldots, m$). Similarly, this process is repeated for the r factor variances, along with their parameters ϕ_{m+j} and σ_{m+j}, with $j = 1, \ldots, R$. The second step entails sampling for each row $\mathbf{\Lambda}_{i,\bullet}$ of the factor loadings matrix, beginning with $\mathbf{\Lambda}_{i,\bullet}|\mathbf{f}, \mathbf{y}_{i,\bullet}, \mathbf{h}_{i,\bullet}$. It is important to note that this step involves estimating m multivariate regressions, each with dimension $\tilde{r}$ and based on T observations, where $\tilde{r}$ represents the number of unconstrained elements in $\mathbf{\Lambda}_{i,\bullet}$.

Following this initial sampling process, new samples are drawn for each element on the main diagonal of $\mathbf{\Lambda}$ using the interweaving approach, either in relation to the equation of state for the factors or concerning latent volatilities. The first case is referred to as shallow interweaving, while the second is termed deep interweaving, as defined in Kastner and Frühwirth-Schnatter (2014). Both strategies aim to accelerate the convergence of Markov chains by re-sampling the parameters conditioned on latent variables of the model in a reparametrized version of the model, with deep interweaving generally preferred in most cases. Finally, the last step of the algorithm involves drawing samples for $\mathbf{f_t}$ from $\mathbf{f_t}|\mathbf{\Lambda}, \mathbf{y_t}, \mathbf{h_t}$ ($t = 1, \ldots, T$), entailing the estimation of an additional T r-dimensional multivariate regressions, each with the number of observations equal to m, representing the number of series in the analysis.

It is worth noting that in step 1, $m + r$ univariate stochastic volatility models are used, which Kastner et al. (2017) combine with two other observation equations:

$$\log(y_{it} - \mathbf{\Lambda}_{i,\bullet}\mathbf{f_t})^2 = h_{it} + \log \varepsilon_{it}^2, \quad i = 1, \ldots, m \tag{23}$$

$$\log f_{jt}^2 = h_{m+j,t} + \log \xi_{jt}^2, \quad j = 1, \ldots, r, \tag{24}$$

and ε_{it} and ξ_{jt} are, respectively, innovation terms associated with $\mathbf{y_t}$ and $\mathbf{f_t}$. The iteration of the algorithm then results in samples from the joint posterior distribution, with the discarding of the burn-in samples in order to avoid any influences from initial values.

3.2. Proposed Formulation

In the context of this study, an alternative formulation has been chosen, where the factor loadings are directly estimated within the latent volatility equations. In general terms, the proposed model, featuring m series and k factors, is represented as follows:

$$\mathbf{r_t} = \mathbf{\Omega}_t \boldsymbol{\varepsilon}_t, \quad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \Sigma_\varepsilon) \tag{25}$$

$$\boldsymbol{\sigma}_t = \boldsymbol{\alpha} + \boldsymbol{\gamma}\mathbf{h_t} + \boldsymbol{\eta}_t, \quad \boldsymbol{\eta}_t \sim N(\mathbf{0}, \Sigma_\eta), \tag{26}$$

where $\boldsymbol{\sigma}_t = (\sigma_{1t}, \ldots, \sigma_{mt})'$ represents a vector containing the m log volatilities, $\boldsymbol{\alpha} = (\alpha_1, \ldots, \alpha_m)'$ denotes a mean parameter, and $\boldsymbol{\gamma}$ is a $m \times k$ matrix of factor loadings. Diagonal covariance matrices Σ_ε and Σ_η are permitted, indicating independent stochastic volatility processes. As for the factor log volatilities, denoted by $\mathbf{h_t} = (h_{1t}, \ldots, h_{kt})'$, we have:

$$\mathbf{h_t} = \mathbf{\Phi}(\mathbf{h_{t-1}}) + \boldsymbol{\xi}_t, \quad \boldsymbol{\xi}_t \sim N(\mathbf{0}, \Sigma_h), \tag{27}$$

where $\mathbf{\Phi} = \text{diag}(\phi_1, \ldots, \phi_k)$ represents a diagonal matrix of persistence parameters, and thus a univariate first-order autoregressive structure for each factor. The relationship between the return equations, represented by the vector $\mathbf{r_t} = (r_{1t}, \ldots, r_{mt})'$, and the volatilities for each asset is expressed through the matrix $\mathbf{\Omega_t}$, which is of size $m \times m$:

$$\mathbf{\Omega_t} = \begin{pmatrix} \exp\{\sigma_{1t}/2\} & 0 & \ldots & 0 \\ 0 & \exp\{\sigma_{2t}/2\} & \ldots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \ldots & \exp\{\sigma_{mt}/2\} \end{pmatrix}. \tag{28}$$

The selected functional forms maintain the desirable properties of the modeling approach using INLA. As argued by Martino (2007), there is a possibility of losing the computational advantage associated with the method in its original form, due to the curse of dimensionality. Therefore, more parsimonious models can help mitigate inefficiencies arising from the dimensionality increase, potentially yielding benefits from increased information compared to univariate modeling. Furthermore, these proposed models can be implemented using the R-INLA package and its existing functionalities.

For the estimation process, we begin with two fundamental identification restrictions: the number of factors must be less than or equal to the number of return series $(k \leq m)$, and for one of these series, the load parameters (γ_{ij}) are fixed on unitary values. Following the implementation procedure of the INLA methodology, the proposed model is reinterpreted as a Gaussian Markov random field model in three stages, as adopted by Martino (2007). In the first stage, a likelihood model is defined:

$$\pi(\mathbf{r}|\boldsymbol{\sigma}, \mathbf{h}, \theta_1) = \prod \pi(\mathbf{r_t}|\boldsymbol{\sigma_t}, \mathbf{h_t}, \theta_1), \tag{29}$$

where θ_1 is a vector of hyperparameters related to volatility.

Next, the latent fields referring to $\boldsymbol{\sigma_t}$ and $\mathbf{h_t}$ are modeled. Then we have:

$$\boldsymbol{\sigma_t}|\mathbf{h_t}, \boldsymbol{\alpha}, \theta_2 \sim N(\boldsymbol{\alpha} + \boldsymbol{\gamma}\mathbf{h_t}, \Sigma_\eta) \tag{30}$$

$$\mathbf{h_t}|\mathbf{h_{t-1}}, \theta_3 \sim N(\mathbf{\Phi h_{t-1}}, \Sigma_h), \tag{31}$$

with θ_2 and θ_3 as hyperparameter vectors associated with their respective covariance matrices (Σ_η and Σ_h), Gaussian prior distributions are assumed for the mean parameters ($\boldsymbol{\alpha}$), centered on zero. Consequently, as demonstrated by Martino (2007), the average volatility can be incorporated into the latent field by calculating the following density:

$$\pi(\boldsymbol{\sigma}, \boldsymbol{\alpha}|\theta_1) = \pi(\boldsymbol{\alpha}) \prod_{t=1}^{T} \pi(\sigma|\mathbf{h_t}, \theta_1) \propto |\mathbf{Q}|^{1/2} \exp\left[-\frac{1}{2}(\boldsymbol{\sigma}', \boldsymbol{\alpha}')\mathbf{Q}(\boldsymbol{\sigma}', \boldsymbol{\alpha}')'\right], \tag{32}$$

where $\mathbf{Q}$ represents the precision matrix, and $(\boldsymbol{\sigma}', \boldsymbol{\alpha}')$ denotes the latent field for volatility. The sparsity of $\mathbf{Q}$ offers computational efficiency, a property explored in greater detail by Rue and Held (2005) and Rue et al. (2009).

In the final stage, a prior $\pi(\boldsymbol{\theta})$ is established for the hyperparameter vectors $\boldsymbol{\theta} = (\theta_1, \theta_2, \theta_3)$. In this context, the study assumes distributions in line with the proposals by Martino (2007) wherever applicable, particularly regarding parameters related to the factor structure. However, for compatibility with the **R-INLA** package, the precision parameter (τ) is handled in terms of its natural logarithm (log-precision), representing a non-informative prior for the precision, and the persistence parameter (ϕ) is transformed using a function defined between -1 and 1, as follows:

$$\ln(\tau_k) \sim logGamma(1, 0.00005), \tag{33}$$

$$\ln\left(\frac{1+\phi_k}{1-\phi_k}\right) \sim N(0, 0.15), \tag{34}$$

the prior distributions of the other parameters are Gaussian. Given the latent field $\mathbf{x} = (\boldsymbol{\sigma}', \boldsymbol{\alpha}')$, the estimation via INLA involves building an approximation for $\pi(x_t|\mathbf{r})$ from $\pi(\boldsymbol{\theta}|\mathbf{r})$ and from $\pi(x_t|\boldsymbol{\theta}, \mathbf{r})$.

To obtain the joint posterior distribution in relation to the hyperparameters, we start from the following Gaussian approximation for the complete conditional density of $\mathbf{x}$:

$$\tilde{\pi}_G(\mathbf{x}|\mathbf{r}, \boldsymbol{\theta}) = \bar{K} \exp\left\{-\frac{1}{2}(\mathbf{x} - \boldsymbol{\nu})'[\mathbf{Q} + \text{diag}(\mathbf{C})](\mathbf{x} - \boldsymbol{\nu})\right\}, \tag{35}$$

$\bar{K}$ is a normalization constant, $\boldsymbol{\nu}$ corresponds to the mode of $\pi(\mathbf{x}|\mathbf{r}, \boldsymbol{\theta})$ and diag$(\mathbf{C})$ is a band matrix with width equal to the number of series m, due to the Markov structure derived from the conditional density, which can be written as:

$$\mathbf{C} = \begin{pmatrix} \mathbf{C_1} & \mathbf{0} & \dots & \mathbf{0} \\ \mathbf{0} & \mathbf{C_2} & \dots & \mathbf{0} \\ \vdots & & \ddots & \vdots \\ \mathbf{0} & \dots & & \mathbf{C_T} \end{pmatrix}, \tag{36}$$

where $\mathbf{C_t}$ contains the 2^d order terms in the Taylor expansion $\sum \log \pi(\mathbf{r_t}|x_t, \theta_1)$ around $\boldsymbol{\nu}$ in the Hessian. In this way, the joint posterior for $\boldsymbol{\theta}$ can be approximated through the following relationship:

$$\tilde{\pi}(\boldsymbol{\theta}|\mathbf{r}) \propto \left.\frac{\pi(\mathbf{r}|\mathbf{x}, \boldsymbol{\theta})\pi(\mathbf{x}|\boldsymbol{\theta})\pi(\boldsymbol{\theta})}{\tilde{\pi}_G(\mathbf{x}|\boldsymbol{\theta}, \mathbf{r})}\right|_{\mathbf{x}=\boldsymbol{\nu}(\boldsymbol{\theta})}. \tag{37}$$

The approximation for $\pi(x_t|\boldsymbol{\theta}, \mathbf{r})$, on the other hand, follows a Gaussian strategy but can take advantage of the results obtained in the evaluation stage $\tilde{\pi}(\boldsymbol{\theta}|\mathbf{r})$. In particular, we take $\tilde{\pi}_G(\mathbf{x}|\boldsymbol{\theta}, \mathbf{r})$ as the mean parameter for the distribution, leaving only the values for the marginal variances, represented by σ^2_G. The calculation is performed according to the recursive methods proposed by Rue and Martino (2007), allowing the approximation to be given by:

$$\tilde{\pi}_G(\mathbf{x}|\boldsymbol{\theta}, \mathbf{r}) = N(x_t; \boldsymbol{\nu}[\boldsymbol{\theta}], \sigma^2_G[\boldsymbol{\theta}]). \tag{38}$$

It is worth pointing out, however, that, according to the authors, this approximation is not accurate in some situations, especially when there are extreme values for $\boldsymbol{\theta}$. Its merit lies in constituting a faster alternative to other more precise ways of performing the calculation, which is especially interesting in the context of models in which dimensionality is a relevant issue.

Finally, $\pi(x_t|\mathbf{r})$ can be approximated, once $\pi(\boldsymbol{\theta}|\mathbf{r})$ and $pi(x_t|\boldsymbol{\theta}, \mathbf{r})$, through a numerical integration of the type:

$$\tilde{\pi}(x_t|\mathbf{r}) = \sum_n \tilde{\pi}(x_t|\boldsymbol{\theta}_n, \mathbf{r})\tilde{\pi}(\boldsymbol{\theta}_n|\mathbf{r})\Delta_n, \quad n \in \{1, \dots, N\}, \tag{39}$$

to be performed on a set of points (*grid*) for $\boldsymbol{\theta}$, with the weights Δ_n being equal to 1 for equidistant points. However, in order to speed up the approximation process, we adopt an approach based on the empirical Bayes procedure, in which only one integration point equivalent to the posterior mode of hyperparameters is used. In general terms, the term $\tilde{\pi}(x_t|\boldsymbol{\theta}_n, \mathbf{r})$ is replaced by $\tilde{\pi}(x_t|\boldsymbol{\theta}_*, \mathbf{r})$, with $\boldsymbol{\theta}_*$ being the mode of $\tilde{\pi}(\boldsymbol{\theta}|\mathbf{r})$. According to Martino (2007), the approach is highly accurate when the distribution of the hyperparameter vector conditional on the log-returns is regular.

3.3. Practical Implementation

From the perspective of practical implementation, the R-INLA package allows more efficient strategies to be defined from a computational point of view with the help of the argument control.inla, which controls the way in which numerical approximations and

integrations are computed. Among the options of interest for the work, the integration approach via the empirical Bayes procedure is achieved through the command control.inla = list(int.strategy = "eb"), while the Gaussian approximation for $\pi(x_t|\boldsymbol{\theta}, \mathbf{r})$ is chosen by the argument strategy = "gaussian". The complete list of arguments considered in the estimation can be seen in the code used:

```
inla(
  formula,
  family = rep("stochvol", m),
  data = data,
  control.inla = list(stupid.search = TRUE,
                      reordering = "metis",
                      strategy = "gaussian",
                      int.strategy = "eb"),
  control.predictor = list(compute = TRUE)
)
```

In addition to the settings already mentioned regarding the strategies for obtaining approximations, one more argument is defined, referring to the reordering algorithm chosen for the precision matrix. The option is to use the METIS implementation[4], which aims to evaluate sparse matrices efficiently, both in terms of storage needs and processing capacity. The fact that METIS reordering is perfectly suited to parallelization stands out, reinforcing the efficiency aspect of matrix operations.

The model is estimated using a multiple likelihood structure, with the same configuration for vectors and indices, regardless of the number of factors taken into account in the analysis. Assuming a general case involving m series, k factors and n observations for each return series we have, as a starting point, the creation of a return matrix of dimensions $(m+1)n \times m$, organized as follows:

$$\mathbf{R} = \begin{pmatrix} NA & NA & \dots & NA \\ \vdots & \vdots & & \vdots \\ NA & NA & \dots & NA \\ r_{1,1} & NA & \dots & NA \\ \vdots & \vdots & & \vdots \\ r_{n,1} & NA & \dots & NA \\ NA & r_{1,2} & \dots & NA \\ \vdots & \vdots & & \vdots \\ NA & r_{n,2} & \dots & NA \\ \vdots & \vdots & & \vdots \\ NA & NA & \dots & r_{1,m} \\ \vdots & \vdots & & \vdots \\ NA & NA & \dots & r_{n,m} \end{pmatrix}, \tag{40}$$

$\mathbf{r}_{t,i}$ refers to the log-return of asset i, for the period t, and null entries in the matrix are denoted by NA. It should also be noted that the first n lines of $\mathbf{R}$ correspond to null values, such that the first entry referring to a return, $\mathbf{r}_{1,1}$, appears in line $n+1$ of the matrix, since in this representation the first n lines of $\mathbf{R}$ represent the unobserved values of the latent factors, treated as missing values and estimated by the Bayesian inference procedure. This formulation implies that the latent factors will be associated with all observed series via the factor loadings structure as copies weighted by the parameters representing the factor loadings of these unobserved variables.

Three other groups of indices are constructed, in parallel, with the purpose of estimating the fixed and random effects of the linear dynamic model. Regarding the fixed effects, we have the following m vectors:

$$\mathbf{a_1} = (\underbrace{NA, \ldots, NA}_{n}, \underbrace{1, \ldots, 1}_{n}, \underbrace{NA, \ldots, NA}_{(m-1)n})' \tag{41}$$

$$\mathbf{a_2} = (\underbrace{NA, \ldots, NA}_{n}, \underbrace{NA, \ldots, NA}_{n}, \underbrace{1, \ldots, 1}_{n}, \underbrace{NA, \ldots, NA}_{(m-2)n})' \tag{42}$$

$$\vdots$$

$$\mathbf{a_m} = (\underbrace{NA, \ldots, NA}_{n}, \underbrace{NA, \ldots, NA}_{(m-1)n}, \underbrace{1, \ldots, 1}_{n})', \tag{43}$$

where the first n elements are left as null values and, for each vector $\mathbf{a_i}$ $(i \in \{1, \ldots, m\})$, a sequence of 1's with dimension n and starting from element number $ni + 1$. Furthermore, other k identical indices associated with the evolution of the factors are added and initialized as follows:

$$\mathbf{f_1} = (1, \ldots, n, \underbrace{NA, \ldots, NA}_{m.n})' \tag{44}$$

$$\vdots$$

$$\mathbf{f_k} = (1, \ldots, n, \underbrace{NA, \ldots, NA}_{m.n})', \tag{45}$$

the first n elements being the sequence from 1 to n, in a structure that is repeated for all $\mathbf{f_j}$, with $j = \{1, \ldots, k\}$. Finally, the random effects with respect to log variances are captured by a set of mk vectors, constructed as:

$$\mathbf{cf_{1,j}} = (\underbrace{NA, \ldots, NA}_{n}, 1, \ldots, n, \underbrace{NA, \ldots, NA}_{(m-1)n})' \tag{46}$$

$$\mathbf{cf_{2,j}} = (\underbrace{NA, \ldots, NA}_{n}, \underbrace{NA, \ldots, NA}_{n}, 1, \ldots, n, \underbrace{NA, \ldots, NA}_{(m-2)n})' \tag{47}$$

$$\vdots$$

$$\mathbf{cf_{m,k}} = (\underbrace{NA, \ldots, NA}_{n}, \underbrace{NA, \ldots, NA}_{(m-1)n}, 1, \ldots, n)', \tag{48}$$

Even if it is true that for the same index i of assets, $\mathbf{cf_{i,j}} = \mathbf{cf_{i,s}}$, $\forall j, s \in \{1, \ldots, k\}$, it is possible to use the same definition of indices for any number $k \leq m$ of factors as long as this last vector structure is created according to the case in which there are as many factors as series in the analysis ($m = k$), taking the number of vectors to m^2. An adjustment would be necessary, however, in the sense of manually assuming that the factors disregarded in the estimation are set to zero, that is, that they have a log-precision set at a high value and their persistence parameter is fixed and equal to zero. See Ruiz-Cárdenas et al. (2012) for details on this structure.

Once the necessary vectors have been constructed, the formula that will enable the estimation of the model is written based on the return matrix and the indices created. In the context of the R-INLA package, terms related to random effects are indicated by the function f(), which takes as arguments the respective index vector, a string of characters informing the model that describes it and additional specifications about parameters. Taking the general case where $m = k$, the formula can be written as:

```
formula = R ~ -1 + a_1 + ... + a_m +
f(f_1, model="ar1",constr=T) + ... +
f(f_m, model="ar1",constr=T) +
f(cf_1.1, copy="f_1", fixed=T, initial=1) +
f(cf_2.1, copy="f_1", fixed=F, hyper=list(...)) + ... +
f(cf_m.1, copy="f_1", fixed=F, hyper=list(...)) + ... +
f(cf_m.m, copy="f_m", fixed=F, hyper=list(...))
```

In this formulation, the models behind the indices of each log-volatility vector are constructed as copies of the structure established for a vector of indices associated with the factors through the attribute copy, which allows for the inclusion of a shared component between more than one linear predictor. Furthermore, it is observed that the models for the vectors of the group $\mathbf{cf_{i,j}}$ have, for each factor j, one of its terms (the first, in this case) presenting fixed hyperparameters, respecting restrictions imposed on factor loadings. For the group of $\mathbf{f_j}$, it is necessary to define the argument "model = "ar1"", which determines a first-order autoregressive structure for the factor log variances.

4. Results

In this section, we analyze the empirical results of estimating the proposed multivariate model, considering specifications with one to four latent factors, using the representation given by the Equations (25) and (26). We compare these results with models estimated using Markov Chain Monte Carlo (MCMC) using the multifactor representation proposed by Kastner et al. (2017) and also with the results of a univariate-based specification. The comparison is based on both in-sample and out-of-sample fit measures, as well as model selection metrics using information criteria.

4.1. Data Description

The empirical part of the work makes use of daily data for stock exchange index quotes[5], over the same time interval. The log-returns for indices representing different markets and from different parts of the globe were selected as variables, namely: Dow Jones Industrial (United States), Ibovespa (Brazil), Índice de Precios y Cotizaciones (Mexico), FTSE 100 (England) and SSE Composite Index (China). In Table 1, these indices are listed, and abbreviations are given to them in order to facilitate their identification and denotation.

Table 1. Market index abbreviations.

Abbreviations	Corresponding Index
DJI	Dow Jones Industrial Average
BOV	Bovespa Index
IPC	S&P/BMV Índice de Precios y Cotizaciones
LSE	FTSE 100
SSE	Shanghai Stock Exchange Composite Index

The analysis covers the daily closing returns of each index from January 2010 to February 2023, representing a sample of 2737 observations for each variable. This time span ensures a sufficient volume of data for the number of series analyzed, facilitating an adequate assessment of the proposed model in terms of computational efficiency. Figure 1 displays the index returns between the specified dates.

The selected indices are widely used in the global financial market and represent various regions worldwide. Additionally, they consider the potential for interaction among them in terms of contagion and spillover effects on the price volatility of traded assets. For indices like Ibovespa and Dow Jones Industrial, Achcar et al. (2012) highlight the significance of their relationship in establishing investment references for emerging markets.

Table 2 provides descriptive statistics associated with the log returns of the selected indices, including mean, median, standard deviation, skewness, kurtosis, and the number

of observations. A Jarque–Bera test is also conducted, with the computed values presented in the table. In all cases, the null hypothesis of Gaussian distribution is rejected at a significance level of 1%.

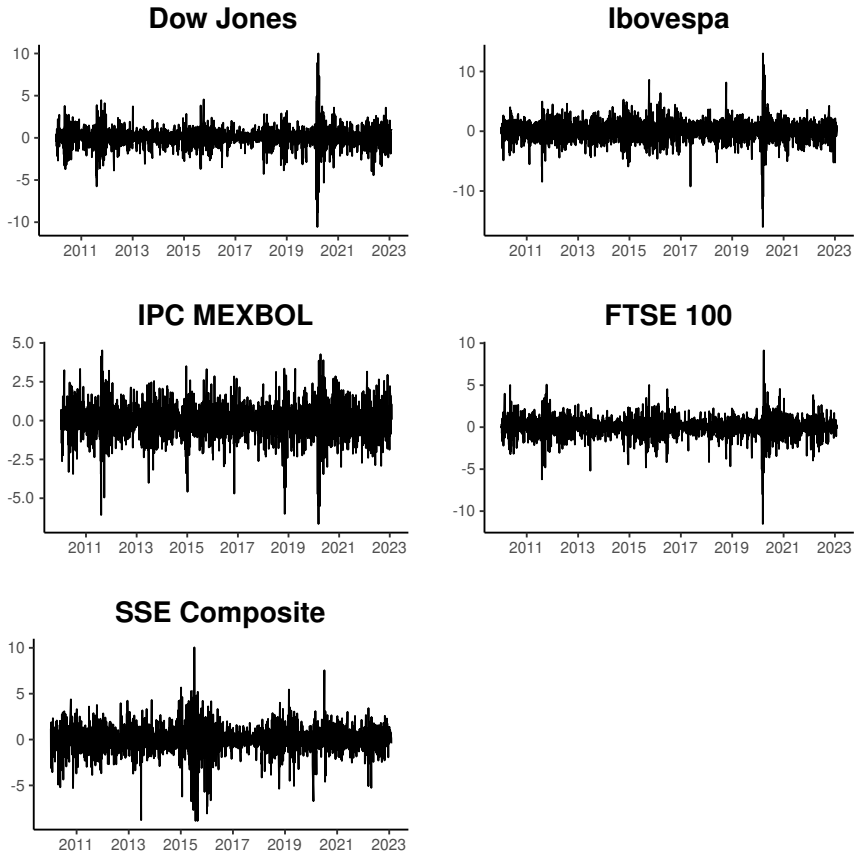

Figure 1. Returns of the analyzed indices.

Table 2. Descriptive statistics of market indices.

	DJI	BOV	IPC	LSE	SSE
Mean	0.042	0.017	0.018	0.012	0.0001
Median	0.074	0.039	0.022	0.055	0.041
Standard Deviation	1.660	0.017	1.045	1.109	1.406
Skewness	−0.509	−0.496	−0.366	−0.606	−0.594
Kurtosis	14.763	12.315	6.343	12.450	9.175
Jarque-Bera	15,894 **	10,005 **	1335 **	10,349 **	4508 **
Sample Size	2737	2737	2737	2737	2737

Note: ** represents p-value < 0.01.

4.2. Empirical Analysis

To analyze the results obtained from the formulation of stochastic volatility models based on the proposed multifactor structure, we examine various specifications for the latent factor structure. We initiate with a structure featuring only one factor and gradually expand the latent factor structure until we reach a model with four factors, using the model represented by Equations (25) and (26) with varying numbers of latent factors. Additionally, we estimate a model based on univariate stochastic volatility (SV) models, which is based on assuming that each observed series only depends on a single latent factor that is unique to this series, but where we estimate these five univariate SV processes jointly using INLA. We compare the in and out of the sample performance of the multivariate factor structure proposed by Kastner et al. (2017) using MCMC methods.

Tables 3–6 provide a summary of the Bayesian posterior distribution of estimated parameters using INLA for the models with one to four latent factors, while Table 7 presents the results for the specification based on univariate SV models. In each table, we present the mean, standard deviation, quantiles at 0.25, 0.5, and 0.975, and the mode of the posterior distribution for each model parameter. Furthermore, we include the marginal likelihood (MLIK), the Deviation Information Criterion (DIC), and the Widely Applicable Information Criterion (WAIC), also known as the Watanabe–Akaike information criterion, for each model. The DIC and WAIC are information criteria frequently used in the context of Bayesian estimation, being a generalization of the Akaike information criterion (AIC). See Ando (2007); Spiegelhalter et al. (2014) for the definition and properties of DIC and Watanabe (2010) for the WAIC.

Table 3. Posterior distribution of estimated parameter—stock index data—Bayesian INLA method—one factor.

		Mean	SD	0.025q	0.5q	0.975q	MODE
h_1	τ_1	0.755	0.058	0.647	0.753	0.876	0.748
	ϕ_1	0.705	0.030	0.644	0.706	0.762	0.706
DJI	α_1	−0.534	0.031	−0.594	−0.534	−0.474	−0.534
	$\gamma_{1,1}$	1.000					
BOV	α_2	0.641	0.029	0.584	0.641	0.699	0.641
	$\gamma_{2,1}$	0.609	0.030	0.549	0.609	0.666	0.611
IPC	α_3	−0.292	0.030	−0.351	−0.292	−0.234	−0.292
	$\gamma_{3,1}$	0.703	0.031	0.641	0.703	0.765	0.703
LSE	α_4	−0.387	0.030	−0.446	−0.387	−0.328	−0.387
	$\gamma_{4,1}$	0.859	0.033	0.794	0.858	0.923	0.858
SSE	α_5	0.337	0.031	0.275	0.337	0.398	0.337
	$\gamma_{5,1}$	0.706	0.036	0.635	0.705	0.777	0.705
MLIK	−20,277.92	DIC	39,490.52	WAIC	39,673.61		

Note: h_i denotes the latent factor i, τ_i denotes the marginal precision for the factor h_i, ϕ_i the persistence parameters for this factor and $\gamma_{k,i}$ denotes the factor loading for the factor i for the series k.

Table 4. Posterior distribution of estimated parameter—stock index data—Bayesian INLA method—two factors.

		Mean	SD	0.025q	0.5q	0.975q	Mode
$\mathbf{h_1}$	τ_1	0.743	0.054	0.641	0.741	0.853	0.740
	ϕ_1	0.688	0.039	0.606	0.690	0.760	0.693
$\mathbf{h_2}$	τ_2	22.435	4.464	15.145	21.933	32.647	20.859
	ϕ_2	0.980	0.004	0.971	0.980	0.987	0.981
DJI	α_1	−0.569	0.031	−0.629	−0.569	−0.509	−0.569
	$\gamma_{1,1}$	1.000					
	$\gamma_{1,2}$	1.000					
BOV	α_2	0.630	0.029	0.572	0.630	0.687	0.630
	$\gamma_{2,1}$	0.615	0.031	0.553	0.615	0.676	0.616
	$\gamma_{2,2}$	0.415	0.145	0.138	0.412	0.708	0.399
IPC	α_3	−0.340	0.030	−0.399	−0.340	−0.281	−0.340
	$\gamma_{3,1}$	0.753	0.035	0.684	0.753	0.823	0.753
	$\gamma_{3,2}$	0.019	0.172	−0.324	0.021	0.353	0.028
LSE	α_4	−0.405	0.030	−0.464	−0.405	−0.346	−0.405
	$\gamma_{4,1}$	0.840	0.043	0.755	0.840	0.926	0.840
	$\gamma_{4,2}$	1.139	0.161	0.824	1.138	1.459	1.133

Table 4. *Cont.*

		Mean	SD	0.025q	0.5q	0.975q	Mode
SSE	α_5	0.256	0.029	0.199	0.256	0.313	0.256
	$\gamma_{5,1}$	0.293	0.050	0.194	0.293	0.392	0.293
	$\gamma_{5,2}$	3.521	0.203	3.126	3.519	3.928	3.511
MLIK	−20,086.77	DIC	39,045.88	WAIC	39,131.66		

Note: h_i denotes the latent factor i, τ_i denotes the marginal precision for the factor h_i, ϕ_i the persistence parameters for this factor and $\gamma_{k,i}$ denotes the factor loading for the factor i for the series k.

Table 5. Posterior distribution of estimated parameter—stock index data—Bayesian INLA method—three factors.

		Mean	SD	0.025q	0.5q	0.975q	Mode
$\mathbf{h_1}$	τ_1	54.085	13.234	32.795	52.497	84.573	49.422
	ϕ_1	0.983	0.005	0.971	0.984	0.992	0.985
$\mathbf{h_2}$	τ_2	2.851	0.203	2.467	2.847	3.264	2.842
	ϕ_2	0.713	0.056	0.593	0.716	0.813	0.721
$\mathbf{h_3}$	τ_3	111.147	43.714	52.792	102.427	221.472	86.726
	ϕ_3	0.969	0.013	0.937	0.971	0.988	0.975
DJI	α_1	−0.585	0.030	−0.644	−0.585	−0.526	−0.585
	$\gamma_{1,1}$	1.000					
	$\gamma_{1,2}$	1.000					
	$\gamma_{1,3}$	1.000					
BOV	α_2	0.592	0.029	0.536	0.592	0.649	0.592
	$\gamma_{2,1}$	0.009	0.478	-0.957	0.018	0.925	0.055
	$\gamma_{2,2}$	1.185	0.100	0.990	1.184	1.385	1.180
	$\gamma_{2,3}$	3.989	0.484	3.094	3.971	4.995	3.888
IPC	α_3	−0.338	0.029	−0.396	−0.338	−0.281	−0.338
	$\gamma_{3,1}$	−0.544	0.376	−1.306	−0.536	0.173	−0.503
	$\gamma_{3,2}$	1.476	0.086	1.310	1.475	1.650	1.469
	$\gamma_{3,3}$	1.282	0.369	0.579	1.275	2.031	1.242
LSE	α_4	−0.413	0.030	−0.471	−0.413	−0.355	−0.413
	$\gamma_{4,1}$	1.440	0.250	0.947	1.440	1.931	1.441
	$\gamma_{4,2}$	1.686	0.096	1.498	1.685	1.877	1.683
	$\gamma_{4,3}$	−0.100	0.363	−0.859	−0.085	0.568	−0.016
SSE	α_5	0.247	0.029	0.191	0.247	0.303	0.247
	$\gamma_{5,1}$	5.178	0.560	4.093	5.172	6.297	5.147
	$\gamma_{5,2}$	0.654	0.133	0.397	0.652	0.921	0.645
	$\gamma_{5,3}$	2.789	0.524	1.780	2.781	3.843	2.747
MLIK	−20,058.90	DIC	38,917.84	WAIC	38,972.25		

Note: h_i denotes the latent factor i, τ_i denotes the marginal precision for the factor h_i, ϕ_i the persistence parameters for this factor and $\gamma_{k,i}$ denotes the factor loading for the factor i for the series k.

We initiated a comprehensive analysis, focusing on the overall model fit, which was measured by marginal likelihood, and model fit was penalized by model complexity, as measured by DIC and WAIC criteria. As expected, the marginal likelihood indicates a better fit for the more complex models, with the model featuring four factors achieving the highest value, calculated as −19,966.49. Additionally, it is worth noting that the model with a single common factor obtains a marginal likelihood of −20,277.92, surpassing that of the model based on univariate stochastic volatility (SV) models, which is calculated as −20,513.56. This indicates that the common factor structure provides an advantage over univariate SV models when considering this model fit criterion.

Table 6. Posterior distribution of estimated parameter—stock index data—Bayesian INLA method—four factors.

		Mean	SD	0.025q	0.5q	0.975q	Mode
$\mathbf{h_1}$	τ_1	12.430	6.381	4.313	11.046	28.697	8.742
	ϕ_1	0.986	0.007	0.969	0.988	0.995	0.990
$\mathbf{h_2}$	τ_2	4.752	1.477	2.501	4.536	8.255	4.133
	ϕ_2	0.944	0.019	0.898	0.947	0.973	0.952
$\mathbf{h_3}$	τ_3	82.015	39.479	31.571	73.512	182.888	59.412
	ϕ_3	0.969	0.010	0.945	0.970	0.984	0.973
$\mathbf{h_4}$	τ_4	2.376	0.467	1.568	2.338	3.399	2.275
	ϕ_4	−0.000	0.058	−0.115	0.000	0.113	0.002
DJI	α_1	−0.585	0.030	−0.644	−0.585	−0.526	−0.585
	$\gamma_{1,1}$	1.000					
	$\gamma_{1,2}$	1.000					
	$\gamma_{1,3}$	1.000					
	$\gamma_{1,4}$	1.000					
BOV	α_2	0.559	0.030	0.501	0.559	0.617	0.559
	$\gamma_{2,1}$	0.127	0.400	−0.652	0.124	0.923	0.113
	$\gamma_{2,2}$	0.920	0.170	0.581	0.921	1.252	0.926
	$\gamma_{2,3}$	3.173	0.520	2.195	3.158	4.239	3.092
	$\gamma_{2,4}$	0.852	0.096	0.661	0.852	1.041	0.854
IPC	α_3	−0.371	0.031	−0.431	−0.371	−0.311	−0.371
	$\gamma_{3,1}$	0.030	0.358	−0.679	0.032	0.732	0.037
	$\gamma_{3,2}$	1.333	0.207	0.923	1.334	1.738	1.337
	$\gamma_{3,3}$	1.330	0.773	−0.136	1.311	2.906	1.230
	$\gamma_{3,4}$	0.960	0.124	0.714	0.961	1.202	0.964
LSE	α_4	−0.416	0.030	−0.475	−0.416	−0.357	−0.416
	$\gamma_{4,1}$	0.994	0.172	0.653	0.995	1.328	1.000
	$\gamma_{4,2}$	1.515	0.140	1.245	1.514	1.795	1.507
	$\gamma_{4,3}$	0.126	0.524	−0.889	0.120	1.173	0.097
	$\gamma_{4,4}$	0.986	0.132	0.723	0.988	1.242	0.994
SSE	α_5	0.198	0.033	0.133	0.198	0.262	0.198
	$\gamma_{5,1}$	2.257	0.541	1.210	2.251	3.338	2.225
	$\gamma_{5,2}$	0.175	0.231	−0.280	0.175	0.631	0.174
	$\gamma_{5,3}$	2.789	0.524	1.780	2.781	3.843	2.747
	$\gamma_{5,4}$	0.778	0.194	0.393	0.778	1.157	0.782
MLIK	−19,966.49	DIC	38,738.78	WAIC	38,766.12		

Note: h_i denotes the latent factor i, τ_i denotes the marginal precision for the factor h_i, ϕ_i the persistence parameters for this factor and $\gamma_{k,i}$ denotes the factor loading for the factor i for the series k.

The DIC and WAIC information criteria also favor more complex models, with the model featuring four latent factors emerging as the best model according to these two criteria. It has DIC and WAIC values of 38,738.78 and 38,766.12, respectively. Even when accounting for greater complexity in the model, models with more latent factors demonstrate a superior fit.

To perform an in-sample disaggregated comparison for each analyzed series and to verify the results of the adjustment of the multifactor models estimated by Markov Chain Monte Carlo (MCMC) using the structure proposed in Kastner et al. (2017), we present Table 8. The table includes the mean error (ME), root mean squared error (RMSE), and mean absolute deviation (MAE) between the absolute value of each return series (taken as a proxy for the true unobserved volatility) and the volatility estimated by the models in our analysis. Bold values denote the best results for each measure in all analyses. We consider the multifactor models with one to four latent factors using INLA estimation, the models with one to four latent factors using MCMC estimation in the context of the

factorial structure proposed in Kastner et al. (2017), and finally, the estimation based on univariate SV models using INLA.

We observe a greater heterogeneity of results. In terms of mean error, the outcomes suggest that specifications with two and three latent factors estimated using INLA perform better for the first four series (DJI, BOV, IPC, and LSE). For SSE, a better result is obtained using specifications with three and four factors estimated using MCMC.

Table 7. Posterior distribution of estimated parameter—stock index data—Bayesian INLA method—univariate based.

		Mean	**SD**	**0.025q**	**0.5q**	**0.975q**	**Mode**
DJI	α_1	−0.432	0.030	−0.491	−0.432	−0.372	−0.432
	τ_1	0.880	0.108	0.686	0.873	1.112	0.859
	ϕ_1	0.940	0.011	0.918	0.941	0.960	0.941
BOV	α_2	0.687	0.029	0.631	0.687	0.744	0.687
	τ_2	2.483	0.366	1.822	2.462	3.259	2.433
	ϕ_2	0.947	0.011	0.922	0.948	0.966	0.950
IPC	α_3	−0.207	0.030	−0.265	−0.207	−0.149	−0.207
	τ_3	1.989	0.298	1.477	1.964	2.648	1.909
	ϕ_3	0.951	0.012	0.925	0.953	0.971	0.955
LSE	α_4	−0.308	0.031	−0.368	−0.308	−0.248	−0.308
	τ_4	1.139	0.128	0.904	1.133	1.409	1.122
	ϕ_4	0.920	0.013	0.892	0.920	0.943	0.921
SSE	α_5	0.267	0.029	0.209	0.267	0.325	0.267
	τ_5	1.429	0.255	0.997	1.406	1.996	1.359
	ϕ_5	0.973	0.007	0.956	0.974	0.985	0.975
MLIK	−20,513.56	DIC	40,074.53	WAIC	40,321.50		

Note: h_i denotes the latent factor i, τ_i denotes the marginal precision for the factor h_i, ϕ_i the persistence parameters for this factor and $\gamma_{k,i}$ denotes the factor loading for the factor i for the series k.

Table 8. In-sample error measures.

			ME	**RMSE**	**MAE**
DJI	INLA	Univariate	0.16179	0.63807	0.47200
		One Factor	0.12713	0.55916	0.41633
		Two Factors	0.11568	0.53136	0.39540
		Three Factors	**0.11171**	0.52809	0.39005
		Four Factors	0.11539	**0.52391**	**0.38893**
	MCMC	One Factor	0.19101	0.65880	0.49337
		Two Factors	0.18940	0.64406	0.47940
		Three Factors	0.18913	0.64344	0.47912
		Four Factors	0.19591	0.63696	0.47890
BOV	INLA	Univariate	0.27936	1.01011	0.78037
		One Factor	0.21693	1.05141	0.77320
		Two Factors	**0.20624**	1.03920	**0.76373**
		Three Factors	0.49439	1.31151	0.90572
		Four Factors	0.32892	1.02122	0.76642
	MCMC	One Factor	0.30308	1.01262	0.78455
		Two Factors	0.30528	**1.00775**	0.78124
		Three Factors	0.30382	1.00492	0.77968
		Four Factors	0.31048	0.97835	0.77221

Table 8. *Cont.*

			ME	RMSE	MAE
IPC	INLA	Univariate	0.18642	0.64376	0.50878
		One Factor	0.13233	0.61804	0.48149
		Two Factors	**0.11898**	**0.58556**	0.45410
		Three Factors	0.24431	0.72142	0.47428
		Four Factors	0.17556	0.59656	**0.43844**
	MCMC	One Factor	0.19705	0.63481	0.50413
		Two Factors	0.20154	0.63808	0.50714
		Three Factors	0.20117	0.63447	0.50472
		Four Factors	0.20492	0.63665	0.50711
LSE	INLA	Univariate	0.16986	0.66251	0.49167
		One Factor	0.12472	0.62571	0.46429
		Two Factors	**0.11625**	**0.61862**	**0.45643**
		Three Factors	0.17439	0.63547	0.46184
		Four Factors	0.22010	0.72893	0.47460
	MCMC	One Factor	0.22477	0.69420	0.52310
		Two Factors	0.22524	0.69771	0.52726
		Three Factors	0.22409	0.69629	0.52645
		Four Factors	0.21856	0.67850	0.51148
SSE	INLA	Univariate	0.26609	0.97774	0.74471
		One Factor	0.82887	2.43031	1.17799
		Two Factors	6.13572	31.41839	6.52991
		Three Factors	0.68351	1.98957	1.02639
		Four Factors	0.29664	0.93203	0.71260
	MCMC	One Factor	0.29663	0.93203	0.71259
		Two Factors	0.29666	0.92906	0.71075
		Three Factors	0.26168	0.82650	0.61935
		Four Factors	**0.26077**	**0.82176**	**0.61520**

Concerning the adjustment results using RMSE, Table 8 reveals that models estimated using INLA with two factors exhibit the best performance for IPC and LSE. For BOV, the estimation by MCMC with two factors is preferable. In the case of DJI, the best result is achieved with the four-factor specification using INLA. For SSE, the four-factor model using MCMC performs the best by this criterion.

The results using the MAE criterion follow a similar pattern. The model with two factors estimated by INLA is selected as the best for BOV and LSE. For DJI and IPC, the four-factor specification using INLA is favored, while the SSE series benefits the most from the four-factor model estimated using MCMC.

To illustrate the fit of the multifactor models, we present Figures 2 and 3, displaying the estimated volatility for the four series using the most general models based on four latent factors. The gray line represents the absolute returns and the red lines the posterior mean of fitted volatilities for INLA and MCMC methods. In general, we can observe that the estimated volatility adequately tracks the movement in absolute returns, both in periods of low and high market volatility.

Finally, we illustrate the temporal evolution of the estimated latent factors in Figure 4, which shows the posterior mean of the latent factors for the model with four latent factors (Table 6), and a credibility interval of 95% constructed using the 2.5% and 97.5% percentiles of the posterior distribution of each latent factor. The black line represents the posterior mean and red lines the 2.5% and 97.5% percentiles of the credibility interval. We can note that the factors capture different persistence patterns, with factors 1, 2 and 3 associated with persistence values close to one in the autoregressive structure of the process, and factor four capturing shocks with almost zero persistence in the conditional volatility series.

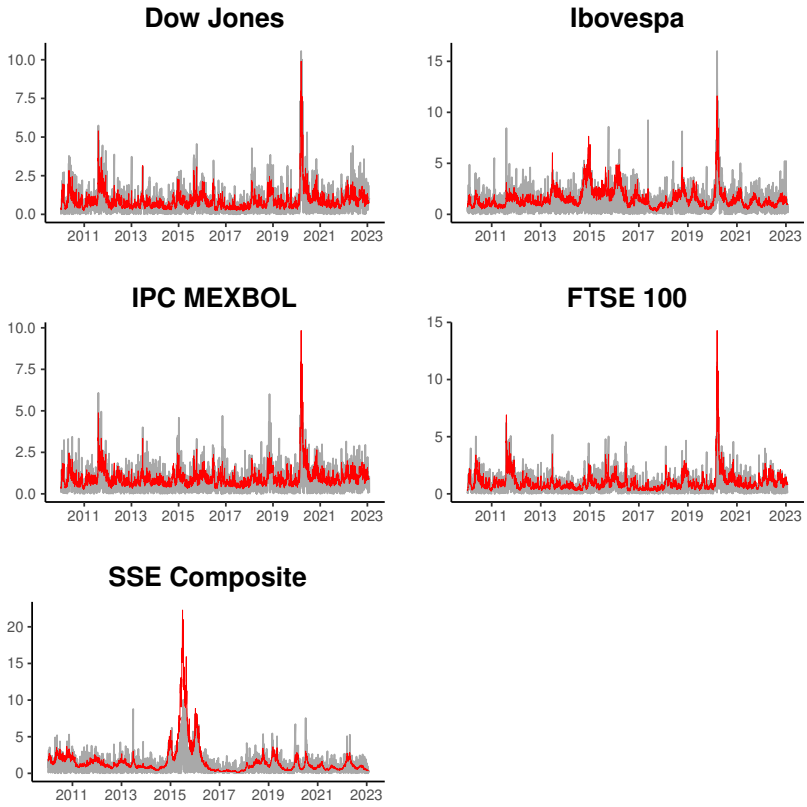

Figure 2. Volatilities estimated via INLA vs. absolute returns.

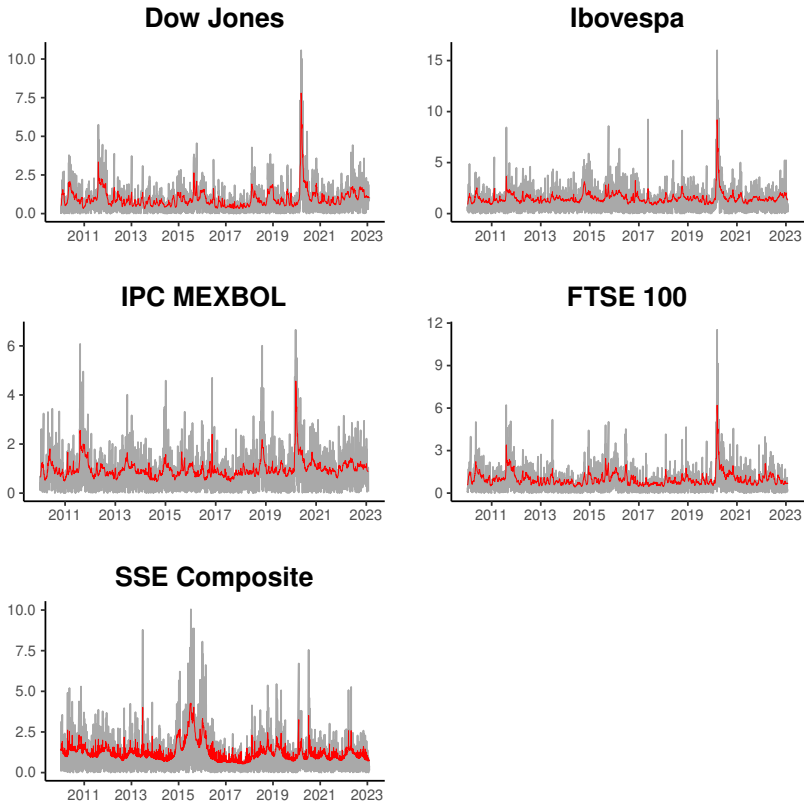

Figure 3. Volatilities estimated via MCMC vs. absolute returns.

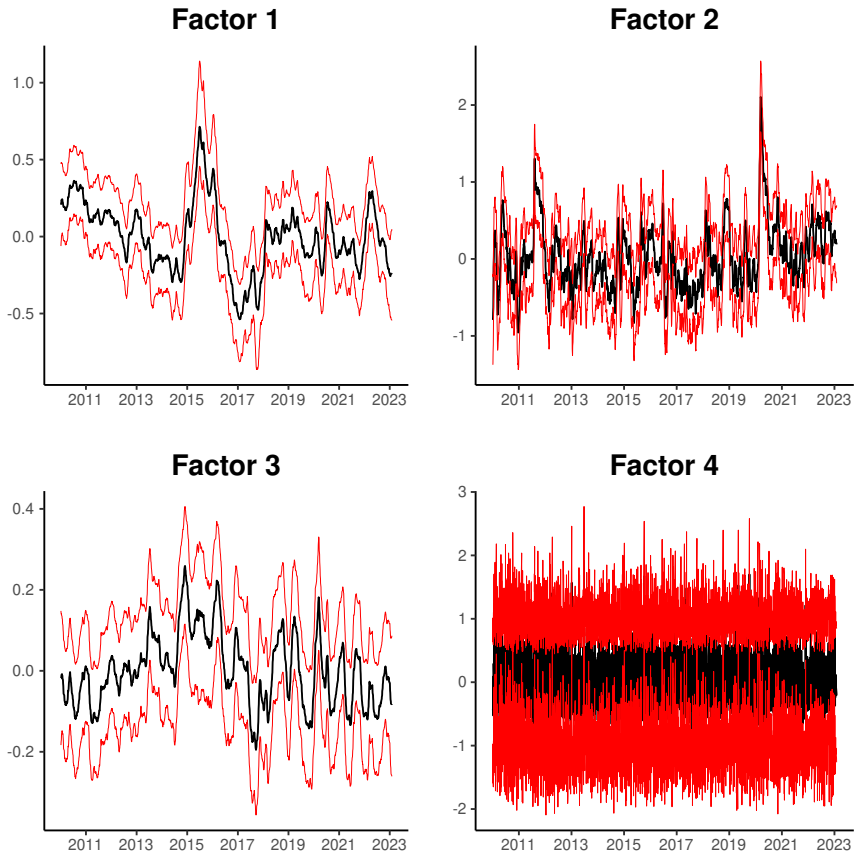

Figure 4. Estimated factors—INLA with four factors.

4.3. Computational Cost

A relevant aspect to consider is the computational cost associated with each estimation. Table 9 displays the elapsed time for estimation (in seconds) for each model for three sample sizes with 500, 1000 and 2714 observations, the last value corresponding to the sample of the analyzed data in the empirical section. Notably, INLA-based estimation demonstrates a significantly reduced computational cost compared to MCMC estimation using the algorithm and specification proposed by Kastner et al. (2017), especially for models with fewer latent factors.

Table 9. Computational cost.

Time in Seconds			
Sample Size	500	1000	2714
Univariate Factors—INLA	3.53	5.14	13.67
One Factor—INLA	2.56	5.59	8.22
Two Factors—INLA	7.78	12.00	41.64
Three Factors—INLA	13.94	33.08	75.58
Four Factors—INLA	19.46	100.47	269.68
One Factor—MCMC	38.12	75.76	197.90
Two Factors—MCMC	170.78	342.01	799.46
Three Factors—MCMC	194.30	391.87	923.77
Four Factors—MCMC	211.82	423.32	1036.11

In the larger sample size, for instance, for single-factor models, MCMC estimation takes 14.47 times longer than INLA estimation in the large sample analysis. However, this cost reduction becomes less pronounced for more complex models. In the case of

models with four factors, the computational cost of MCMC estimation is approximately 3.84 times higher.

Comparing the sample sizes the computational cost of the INLA method appears to grow faster with an increasing number of factors compared to the MCMC-based method, but continues to be significantly lower in all analyses. It is essential to note that evaluating the effective cost of MCMC estimation in real situations can be more complex, as it depends on the convergence properties of the chains. In this experiment, we used chains of size 40,000, discarding the first 3500 samples. Still, for other models and sample sizes, the choice of the number of MCMC samples may differ.

The elapsed times were measured on a server with the following configuration: an Intel(R) Xeon(R) W-2265 processor with 24 threads running at 3.50 GHz and 128 GB of RAM memory. We used INLA version 23.09.09 (built on 9 September 2023). An important note is that we did not utilize the PARDISO sparse linear algebra library to speed up INLA calculations due to licensing restrictions at the time of writing the article.

4.4. Out-of-Sample Volatility Forecasting

To assess the predictive performance of the factor models of stochastic volatility, we conducted an out-of-sample forecast analysis, comparing the predictions of the factor models with one to four factors and the univariate stochastic volatility (SV)-based model estimated using INLA.

In this experiment, we generated out-of-sample predictions, forecasting one and five steps ahead, for the last 22 observations in our sample. We did this by employing a rolling sample approach, incrementing one observation at each step in the estimation sample. As with the in-sample analysis, we report the mean error (ME), root mean squared error (RMSE), and mean absolute error (MAE) between the predicted volatility values generated by each model and the absolute values of each series, serving as a proxy for true unobserved volatility, comparing the INLA and MCMC approaches.

Tables 10 and 11 provide a summary of the predictive results obtained in this analysis. For the one-step-ahead forecast (Table 10), the results reveal a heterogeneous outcome regarding the best model for this horizon. For DJI, the three-factor model is favored, and for SSE, the four-factor model is chosen, both using the INLA approach. Remarkably, all three criteria consistently select the best model for each series. The specification based on univariate SV using INLA is chosen for IPC and LSE, while for BOV, the one-factor model estimated using MCMC is selected.

In the case of the five-step-ahead forecast horizon, as presented in Table 11, the results suggest a better performance of more complex models, with the exception of the LSE and BOV series. Once again, the three criteria are consistent across all series, except for BOV. The INLA-based four-factor model is selected for DJI, and SSE, while the three-factor model using INLA is chosen for CPI. Conversely, the best model for LSE is based on a specification relying on univariate SV models using INLA, and for BOV the univariate models using MCMC are selected by the ME and MAE criteria, and the one-factor model using MCMC by the MSE criterion.

We finish the analysis of the predictive performance with a comparison of the predictive result for the multi-step forecasts for the last 22 observations in the sample without the use of rolling samples, comparing the predictive densities using the Log-Predictive Likelihood (Gelman et al. 2014), again comparing the factor models proposed in the article with the formulation proposed by Kastner et al. (2017). In this analysis, we compared the predictive density for predictions 1, 5, 10 and 22 steps ahead, through the construction of the predictive log-likelihood obtained using the model predictions in a multivariate Gaussian distribution function. The advantages of this method are that it does not depend on any proxy for unobserved variance, and it allows evaluating predictions for all series simultaneously. Table 12 show the results of this analysis.

Table 10. Out-of-sample error measures—One step ahead.

			ME	RMSE	MAE
DJI	INLA	Univariate	0.23196	0.52897	0.42865
		One Factor	0.23777	0.53159	0.43246
		Two Factors	0.23146	0.52808	0.42822
		Three Factors	**−0.02255**	**0.47559**	**0.37538**
		Four Factors	0.12455	0.49130	0.38182
	MCMC	Univariate	0.42554	0.63844	0.56427
		One Factor	0.35854	0.59616	0.51447
		Two Factors	0.42031	0.63643	0.55885
		Three Factors	0.42035	0.63613	0.56240
		Four Factors	0.41900	0.63277	0.55788
BOV	INLA	Univariate	0.59688	1.21295	0.98283
		One Factor	0.44138	1.14441	0.89492
		Two Factors	0.46336	1.15300	0.90732
		Three Factors	0.70404	1.27005	1.04811
		Four Factors	0.51051	1.16925	0.93029
	MCMC	Univariate	0.15588	1.06588	0.76938
		One Factor	**0.08921**	**1.05869**	**0.74359**
		Two Factors	0.15097	1.06256	0.76384
		Three Factors	0.15101	1.06645	0.76832
		Four Factors	0.14967	1.06514	0.76602
IPC	INLA	Univariate	**0.15683**	**0.61506**	**0.50386**
		One Factor	0.20836	0.63027	0.51995
		Two Factors	0.25486	0.64773	0.54613
		Three Factors	0.21796	0.63278	0.52387
		Four Factors	0.38826	0.71054	0.62008
	MCMC	Univariate	0.46614	0.75602	0.66833
		One Factor	0.40088	0.71787	0.62867
		Two Factors	0.46265	0.75619	0.66960
		Three Factors	0.46269	0.75148	0.66482
		Four Factors	0.46134	0.75479	0.66735
LSE	INLA	Univariate	**0.24449**	**0.47201**	**0.42422**
		One Factor	0.58243	0.70863	0.65335
		Two Factors	0.56198	0.69252	0.63858
		Three Factors	0.42964	0.58955	0.54046
		Four Factors	0.63417	0.75155	0.69459
	MCMC	Univariate	0.79527	0.89223	0.83042
		One Factor	0.72887	0.83358	0.77455
		Two Factors	0.79064	0.88803	0.82540
		Three Factors	0.79068	0.88933	0.82655
		Four Factors	0.78933	0.88529	0.82328
SSE	INLA	Univariate	0.32571	0.60216	0.53065
		One Factor	0.84922	0.98873	0.89373
		Two Factors	0.22027	0.54569	0.48797
		Three Factors	−0.35623	0.61848	0.43029
		Four Factors	**0.01095**	**0.50061**	**0.41794**
	MCMC	Univariate	0.71897	0.87956	0.78779
		One Factor	0.65316	0.82586	0.73876
		Two Factors	0.71493	0.87635	0.78583
		Three Factors	0.71497	0.87491	0.78371
		Four Factors	0.71362	0.87467	0.78329

Table 11. Out-of-sample error measures—Five steps ahead.

			ME	RMSE	MAE
DJI	INLA	Univariate	0.15225	0.50562	0.38903
		One Factor	0.51637	0.70659	0.64482
		Two Factors	0.44178	0.65637	0.58746
		Three Factors	−0.10767	0.49464	0.37931
		Four Factors	**−0.06229**	**0.48575**	**0.37375**
	MCMC	Univariate	0.35322	0.59870	0.51845
		One Factor	0.36097	0.60159	0.52346
		Two Factors	0.45301	0.66791	0.60115
		Three Factors	0.45485	0.65998	0.59577
		Four Factors	0.45156	0.66221	0.59627
BOV	INLA	Univariate	0.56319	1.26126	1.02468
		One Factor	0.55355	1.25678	1.01814
		Two Factors	0.54222	1.25079	1.00996
		Three Factors	0.46032	1.21843	0.95592
		Four Factors	0.21479	1.14993	0.81975
	MCMC	Univariate	**0.06254**	1.13142	**0.75413**
		One Factor	0.06939	**1.13082**	0.75775
		Two Factors	0.16142	1.13898	0.79277
		Three Factors	0.16326	1.14303	0.80124
		Four Factors	0.15997	1.13582	0.79787
IPC	INLA	Univariate	0.15762	0.66886	0.54105
		One Factor	0.25359	0.69787	0.58380
		Two Factors	0.27772	0.70832	0.59733
		Three Factors	**0.03243**	**0.64996**	**0.52217**
		Four Factors	0.11032	0.66077	0.52900
	MCMC	Univariate	0.32007	0.72419	0.61349
		One Factor	0.32909	0.72929	0.61745
		Two Factors	0.42113	0.77369	0.66358
		Three Factors	0.42297	0.77900	0.66648
		Four Factors	0.41968	0.77720	0.66800
LSE	INLA	Univariate	**0.22963**	**0.36349**	**0.32397**
		One Factor	0.84995	0.89541	0.84995
		Two Factors	0.78686	0.83671	0.78686
		Three Factors	0.37292	0.46756	0.41977
		Four Factors	0.50862	0.58195	0.52605
	MCMC	Univariate	0.80137	0.84953	0.80137
		One Factor	0.80892	0.85672	0.80892
		Two Factors	0.90096	0.94456	0.90096
		Three Factors	0.90279	0.94593	0.90279
		Four Factors	0.89950	0.94130	0.89950
SSE	INLA	Univariate	0.31107	0.52824	0.47098
		One Factor	1.05909	1.14185	1.05909
		Two Factors	0.42805	0.62888	0.55599
		Three Factors	−0.36826	0.56374	0.41105
		Four Factors	**−0.06832**	**0.43178**	**0.35291**
	MCMC	Univariate	0.69274	0.81373	0.72497
		One Factor	0.70147	0.82115	0.73149
		Two Factors	0.79351	0.89953	0.79937
		Three Factors	0.79534	0.90402	0.80667
		Four Factors	0.79205	0.89677	0.79682

Table 12. Out-of-sample forecasting—Log-predictive likelihood.

Horizon	1-Step	5-Steps	10-Steps	22-Steps
One Factor—INLA	−**8.124187**	−7.289769	−8.864761	−6.430457
Two Factors—INLA	−8.312072	−8.588173	−9.940037	−7.359121
Three Factors—INLA	−8.263323	−6.877244	−**8.448241**	−6.013806
Four Factors—INLA	−8.274462	−7.542268	−9.075323	−7.441591
One Factor—MCMC	−8.432325	−6.226095	−8.573205	−5.760034
Two Factors—MCMC	−8.496381	−6.214516	−8.658612	−**5.713363**
Three Factors—MCMC	−8.488097	−6.254625	−8.613700	−5.747760
Four Factors—MCMC	−8.504874	−**6.194094**	−8.614639	−5.777989

In this analysis, we can observe that the predictive results were balanced between the two types of models, with the best result for 1-step and 10-steps ahead using INLA with one and three factors, and using MCMC for 5 and 22 steps ahead.

5. Conclusions

We introduce a multivariate extension to the class of stochastic volatility (SV) models estimated via Integrated Nested Laplace Approximations (INLA), incorporating a multifactor structure that enables a more parsimonious representation of this class of models. Our investigation focuses on the volatility dynamics of stock indexes.

We estimated the model using the log-returns of the Dow Jones Industrial, Ibovespa, Índice de Precios y Cotizaciones, FTSE 100, and SSE Composite Indexes. Subsequently, we assessed its performance in terms of accuracy and computational efficiency. Compared to the Markov Chain Monte Carlo (MCMC) approach, our results indicate that estimating multivariate SV models with INLA is feasible, with little to no compromise in terms of model fit quality. Additionally, there are significant time savings when using INLA to estimate the model parameters.

Regarding in-sample accuracy, our proposed formulation performed as well as or very close to its MCMC counterpart for nearly all assets and across various metrics. While there was no substantial difference in the numbers overall, it is worth noting that the MCMC estimation consistently outperformed the INLA-based approach for the Shanghai Chinese Stock Exchange Index (SSE). Nevertheless, INLA demonstrated a substantial reduction in estimation time, up to 14.47 times less than the MCMC alternative. This efficiency gap tends to decrease as more factors are introduced into the analysis.

We also conducted out-of-sample forecasting to evaluate the predictive performance of the factor stochastic volatility models. For the five-step-ahead forecast horizon, the more complex models (three and four-factor) consistently outperformed univariate models for each series, also estimated using INLA. However, for the one-step-ahead horizon, the univariate SV specification was preferred for the Índice de Precios y Cotizaciones (IPC) and FTSE 100 (LSE).

In conclusion, our work opens the door to further generalizations of the model, which could provide a deeper understanding of volatility patterns across a broader range of financial assets. Potential avenues include conducting Value at Risk (VaR) analyses and introducing additional specifications to better model the characteristics often attributed to financial time series, such as the presence of leverage effects and long memory processes. In the case of long memory processes, the analysis of intraday data may be particularly promising, as the large volume of observations should favor the use of methodologies like INLA over traditional MCMC-based approaches.

An interesting extension of our analysis would be a comparison of computational performance and predictive properties compared to estimations using Variational Bayes (Gunawan et al. 2021; Tan and Nott 2018), which also allows relevant computational gains in the Bayesian estimation of complex models. It is noteworthy that the recent implementation of the INLA methodology enables the incorporation of a Variational Bayes correction to the Gaussian approximation utilized in INLA for certain model classes, as discussed in Van

Niekerk et al. (2023), thereby integrating the computational benefits of this method class into the use of Laplace approximations in Bayesian model estimation.

Author Contributions: Conceptualization, J.P.C.d.S.M.N. and M.L.; methodology, J.P.C.d.S.M.N. and M.L.; Formal analysis, J.P.C.d.S.M.N. and M.L.; investigation, J.P.C.d.S.M.N. and M.L.; writing—original draft preparation, J.P.C.d.S.M.N. and M.L.; writing—review and editing, J.P.C.d.S.M.N. and M.L. All authors have read and agreed to the published version of the manuscript.

Funding: The authors acknowledge funding from Capes, CNPq (310646/2021-9) and FAPESP (2023/02538-0).

Data Availability Statement: All data taken from https://www.investing.com/indices, accessed on 10 March 2023.

Acknowledgments: We are grateful for the comments and criticisms of the Academic Editor and three anonymous reviewers.

Conflicts of Interest: The authors declare no conflict of interest.

Notes

1 https://www.r-inla.org/, accessed on 31 December 2023.

2 Hosszejni and Kastner (2021) decompose the covariance matrix so that $\Sigma_t = \Lambda \tilde{\Sigma}_t \Lambda' + \bar{\Sigma}_t$.

3 The Ancillarity-Sufficiency Interweaving Strategy is extensively discussed by Kastner and Frühwirth-Schnatter (2014).

4 See http://glaros.dtc.umn.edu/gkhome/metis/metis/overview, accessed on 31 December 2023.

5 All data taken from https://www.investing.com/indices, accessed on 10 March 2023.

References

Achcar, Jorge Alberto, Edilberto Cepeda-Cuervo, and Milton Barossi-Filho. 2012. Multivariate volatility models: An application to Ibovespa and Dow Jones Industrial. *Cuadernos de Economía* 31: 301–20.

Andersen, Torben, Bent Sorensen, and Hyung-Jin Chung. 1999. Efficient method of moments estimation of a stochastic volatility model: A Monte Carlo study. *Journal of Econometrics* 91: 61–87. [CrossRef]

Ando, Tomohiro. 2007. Bayesian predictive information criterion for the evaluation of hierarchical bayesian and empirical bayes models. *Biometrika* 94: 443–58. [CrossRef]

Asai, Manabu, Michael McAleer, and Jun Yu. 2006. Multivariate stochastic volatility: A review. *Econometric Reviews* 25: 145–75. [CrossRef]

Black, Fischer. 1976. Studies of stock market volatility changes. In *Proceeding of the 1976 Meetings of the Business and Economics Statistics Section*. Washington, DC: American Statistical Association, pp. 177–81.

Bollerslev, Tim. 1986. Generalized Autoregressive Conditional Heteroskedasticity. *Journal of Econometrics* 31: 307–27. [CrossRef]

Bollerslev, Tim. 1990. Modelling the coherence in short-run nominal exchange rates: A multivariate generalized ARCH model. *The Review of Economics and Statistics* 72: 498–505. [CrossRef]

Chaim, Pedro, and Márcio P. Laurini. 2019. Essays in Financial Econometrics. Ph.D. dissertation, Faculdade de Economia, Administração e Contabilidade de Ribeirão Preto, Ribeirão Preto, Brazil.

Danıelssonn, Jó. 1998. Multivariate stochastic volatility models: Estimation and a comparison with VGARCH models. *Journal of Empirical Finance* 5: 155–73. [CrossRef]

Ehlers, Ricardo, and Mauricio Zevallos. 2015. Bayesian estimation and prediction of stochastic volatility models via INLA. *Communications in Statistics—Simulation and Computation* 44: 683–93. [CrossRef]

Engle, Robert. 1982. Autoregressive Conditional Heteroskedasticity with estimates of the variance of United Kingdom inflation. *Econometrica* 50: 987–1007. [CrossRef]

Frühwirth-Schnatter, Sylvia, Darjus Hosszejni, and Hedibert Freitas Lopes. 2023. When it counts—Econometric identification of the basic factor model based on glt structures. *Econometrics* 11: 26. [CrossRef]

Gelman, Andrew, Jungyoon Hwang, and Aki Vehtari. 2014. Understanding predictive information criteria for Bayesian models. *Statistics and Computing* 24: 997–1016. [CrossRef]

Gunawan, David, Robert Kohn, and David Nott. 2021. Variational Bayes approximation of factor stochastic volatility models. *International Journal of Forecasting* 37: 1355–75. [CrossRef]

Harvey, Andrew. 1989. *Forecasting, Structural Time Series and the Kalman Filter*. Cambridge: Cambridge University Press.

Harvey, Andrew, and Neil Shephard. 1996. Estimation of an asymmetric stochastic volatility model for asset returns. *Journal of Business & Economic Statistics* 14: 429–34.

Harvey, Andrew, Esther Ruiz, and Neil Shephard. 1994. Multivariate stochastic variance models. *The Review of Economic Studies* 61: 247–64. [CrossRef]

Hosszejni, Darjus, and Gregor Kastner. 2021. Modeling Univariate and Multivariate Stochastic Volatility in R with stochvol and factorstochvol. *Journal of Statistical Software* 100: 1–34. [CrossRef]
Jacquier, Eric, Nicholas Polson, and Peter Rossi. 1995. *Models and Priors for Multivariate Stochastic Volatility*. Technical Report. Montréal: Centre Interuniversitaire de Recherche en Analyse des Organisations (CIRANO).
Kastner, Gregor, and Sylvia Frühwirth-Schnatter. 2014. Ancillarity-sufficiency interweaving strategy (ASIS) for boosting MCMC estimation of stochastic volatility models. *Computational Statistics & Data Analysis* 76: 408–23.
Kastner, Gregor, Sylvia Frühwirth-Schnatter, and Hedibert Lopes. 2017. Efficient Bayesian inference for multivariate factor stochastic volatility models. *Journal of Computational and Graphical Statistics* 26: 905–17. [CrossRef]
Kim, Sangjoon, and Neil Shephard. 1998. Stochastic volatility: Likelihood inference and comparison with ARCH models. *The Review of Economic Studies* 65: 361–93. [CrossRef]
Martino, Sara. 2007. *Approximate Bayesian Inference for Multivariate Stochastic Volatility Models*. Technical Report. Trondheim: Department of Mathematical Sciences, Norwegian University of Science and Technology.
Martino, Sara, Kjerti Aas, Ola Lindqvist, Linda Neef, and Havard Rue. 2011. Estimating stochastic volatility models using integrated nested Laplace approximations. *European Journal of Finance* 17: 487–503. [CrossRef]
Martins, Thiago, Daniel Simpson, Finn Lindgren, and Håvard Rue. 2013. Bayesian computing with INLA: New features. *Computational Statistics & Data Analysis* 67: 68–83.
Nelson, Daniel. 1988. *The Time Series Behavior of Stock Market Volatility and Returns*. Cambridge, MA: Doutorado em economia, Massachusetts Institute of Technology.
Quintana, Jose, and Mike West. 1987. An analysis of international exchange rates using multivariate DLM's. *Journal of the Royal Statistical Society: Series D (The Statistician)* 36: 275–81. [CrossRef]
Rajaratnam, Bala, and Doug Sparks. 2015. MCMC-based inference in the era of big data: A fundamental analysis of the convergence complexity of high-dimensional chains. *arXiv* arXiv:1508.00947.
Ravishanker, Nalini, Balaji Raman, and Refik Soyer. 2022. *Dynamic Time Series Models Using R-INLA: An Applied Perspective*. Boca Raton: CRC Press.
Rue, Havard, and Leonhard Held. 2005. *Gaussian Markov Random Fields: Theory and Applications*. Boca Raton: CRC Press.
Rue, Havard, and Sara Martino. 2007. Approximate Bayesian inference for hierarchical Gaussian Markov random field models. *Journal of Statistical Planning and Inference* 137: 3177–92. [CrossRef]
Rue, Havard, Sara Martino, and Nicholas Chopin. 2009. Approximate Bayesian inference for latent gaussian models by using integrated nested Laplace approximations. *Journal of the Royal Statistical Society* 71: 319–92. [CrossRef]
Ruiz-Cárdenas, Ramiro, Elias Krainski, and Håvard Rue. 2012. Direct fitting of dynamic models using integrated nested laplace approximations—INLA. *Computational Statistics & Data Analysis* 56: 1808–28.
Sandmann, Gleb, and Siem Jan Koopman. 1998. Estimation of stochastic volatility models via Monte Carlo maximum likelihood. *Journal of Econometrics* 87: 271–301. [CrossRef]
Shapovalova, Yuliya. 2021. "Exact" and approximate methods for Bayesian inference: Stochastic volatility case study. *Entropy* 23: 466. [CrossRef] [PubMed]
Spiegelhalter, David J., Nicola G. Best, Bradley P. Carlin, and Angelika van der Linde. 2014. The deviance information criterion: 12 years on (with discussion). *Journal of the Royal Statistical Society, Series B* 76: 485–93. [CrossRef]
Tan, Linda S. L., and David J. Nott. 2018. Gaussian variational approximation with sparse precision matrices. *Statistics and Computing* 28: 259–75. [CrossRef]
Taylor, Stephen. 1986. *Modelling Financial Time Series*. Hoboken: John Wiley & Sons.
Van Niekerk, Janet, Elias Krainski, Denis Rustand, and Håvard Rue. 2023. A new avenue for bayesian inference with INLA. *Computational Statistics & Data Analysis* 181: 107692.
Watanabe, Sumio. 2010. Asymptotic equivalence of bayes cross validation and widely applicable information criterion in singular learning theory. *Journal of Machine Learning Research* 11: 3571–94.
Yu, Jun, and Renate Meyer. 2006. Multivariate stochastic volatility models: Bayesian estimation and model comparison. *Econometric Reviews* 25: 361–84. [CrossRef]