Journal of Mathematics Research; Vol. 15, No. 1; February 2023
ISSN 1916-9795 E-ISSN 1916-9809
Published by Canadian Center of Science and Education

From Multidimensional Ornstein - Uhlenbeck Process to Bayesian Vector Autoregressive Process

Lewis N.K. Mambo

Correspondence: Lewis N.K. Mambo, University of Kinshasa and Central Bank of Congo, D.R.C, Congo

Received: October 20, 2022 Accepted: January 28, 2023 Online Published: February 1, 2023

doi:10.5539/jmr.v15n1p32 URL: https://doi.org/10.5539/jmr.v15n1p32

Abstract

The main purpose of this paper is to make the connexion between stochastic analysis, the Bayesian Statistics, and time series analysis for policy analysis. This approach solves the problem of mathematical modelling - the presence of uncertainties in the models and parameters - that reduces the policy analysis and forecasting effectiveness. By using the multiple Itô integral, the multidimensional Ornstein - Uhlenbeck process can be written as a Vector Autoregressive with lag 1 (VAR(1)) that is the generalization of Vector Autoregressive process. The limit of this approach is in fact it requires the strong foundations of stochastic analysis, the Bayesian Statistics, and time series analysis.

Keyword: Bayesian parameter estimation, Vector Autoregressive process, Multiple Itô integrals, Stochastic differential equation, Policy analysis,Uncertainties, Democratic Republic of the Congo

MSC2020: 60G15, 62C10, 62H12, 62F05.

1. Introduction

In the last six decades and in mathematical modelling, the stochastic analysis and time series analysis are more used by many scientists and policy-makers. Indeed, the stochastic differential equations(SDEs) have plaid a major role in the fields of Finance and Financial econometrics. The Ornstein - Uhlenbeck process is a stochastic process that represents the real - world in the stochastic differential Equation.

The Vector Autoregressive models are one of the most successful statistical modelling ideas that have came up in the last forty years. They take into account the dynamic behaviour of the real-world. The VAR model is the multivariate of the autoregressive model. The use of Bayesian estimation methods makes the VAR models generic enough to handle a variety of complex real - world time series Hamilton (1994), Lutkepohl and Kratzig (2004),Lutkepohl (2005), Sim (1990).

Bayesian vector autoregressive (BVARs) models are standard multivariate autoregressive models estimated with Bayesian methods and are used in empirical macroeconomic analysis and finance for the forecasting, structural analysis, and scenario analysis. The Bayesian VAR is the application of Bayes theorem that was laid down in a revolutionary paper written by British mathematician and Reverend Thomas Bayes (1702 - 1761), which appeared in print in 1763 but was not acknowledged for its significance Berger (1985), Kadiyala and Karlsson (1997), Amemiya (1985). Bayesian statistics provides a rational theory of personal beliefs compounded with real world data in the context of uncertainty. In the last three decades, Bayesian Statistics has emerged as one of the leading paradigms in which all of this can be done in a unified fashion. There has been tremendous development in Bayesian theory, methodology, Chin and Li (2019), computation and applications in the past several years Robert (2007),Koop (2003),Zellner (1971), Kelly and Smith (2011), Gelman et al. (2004), Geweke (2005).

These problems of uncertainties modelling are resolved by the use of the stochastic differential equations models and the Bayesian method. These two tools are less used and understood because of their more complex mathematical foundations and implementation Robert (2007),Prakasa Rao (2010), Oksendal (2000),

Basawa et al. (2001),Karatzas and E.Shreve (1988), Kutoyants (1998), Kutoyants (2004), Kushner and Yin (2003), Kushner and P.Dupuis (2001).

For scientists with little or no formal statistical background, Bayesian methods are being discovered as the only viable method for approaching their problems. For many of them, statistics has become synonymous with Bayesian statistics O'Hagan (2000), O'Hagan and West (2010). The Bayesian Vector Autoregressive models are one of the most successful statistical modelling ideas that have came up in the last four decades. The use of Bayesian methods makes the models generic enough to handle a variety of complex real - world time series.

The purpose of Bayesian inference is to provide a mathematical machinery that can be used for modelling systems, where the uncertainties of the system are taken into account and the decisions are made according to rational principles Gelman et al. (2004), Robert (2007).

The rest of this paper is organized as follows. Section 2 gives the preliminaries and notations that constitute the mathematical foundations of the stochastic differential equations and time series analysis. The Section 3 presents the multidimensional Ornstein - Uhlenbeck process. The Section 4 gives the Bayesian vector autoregressive models. And the last section gives the results of empirical analysis from the Democratic Republic of the Congo's data.

2. Preliminaries and Notations

The following notation is used in relation to a matrix A: transpose $A^{'}$, inverse A^{-1}, trace tr(A), and determinant $\mid A \mid$, estimated $\hat{A}$ respectively. For an $n \times m$ matrix A of full column rank (n ¿ m), an orthogonal complement is denoted by $A^{\perp}$. The zero matrix is the orthogonal complement of a non-singular square matrix, and an identity matrix of suitable dimension is the orthogonal complement of a zero matrix. The $vec(A)$ denotes the column-stacking vector of matrix A, $\otimes$ signifies the Kronecker product, and I_T is an $T \times T$ identity matrix. Also $\mathbb{C}$ is a complex set and i means imaginary , $i = \sqrt{-1}$.

Definition 2.1. *(Brownian motion) A Brownian motion is a continuous, adapted process*$B = \{B_t, \mathcal{F}_t : 0 \leq s < \infty$*, defined on some probability space* $(\Omega, \mathcal{F}, P)$*, with the properties that* $B_0 = 0$ *a.s. and for* $0 \leq s < \infty$*, the increment* $B_t - B_s$ *is independent of* $\mathcal{F}_f$ *and is normally distributed with mean zero and variable* $t - s$*.*

Definition 2.2. *(Brownian motion with respect to a filtration) A vectorial* $(d - dimensional)$ *Brownian motion on* $\boldsymbol{T}$ *with respect to a filtration* $\mathbb{F} = (\mathcal{F}_t)_{t\in \boldsymbol{T}}$ *such that (i)* $W_0 = 0$*; (ii)For all* $0 \leq s < t$ *in* $\boldsymbol{T}$*, the increment* $W_t - W_s$ *is independent of* $\mathcal{F}_s$ *and follows a centred Gaussian distribution with variance - covariance matrix* $(t - s)I_d$*.*

One can state a classical property of Brownian motion as follows.

Proposition 2.0.1. *Let* $(W_t)_{t\in \boldsymbol{T}}$ *be a Brownian motion with respect to* $(\mathcal{F}_t)_{t\in \boldsymbol{T}}$*, as follows (1) symmetry:* $(W_t)_{t\in \boldsymbol{T}}$ *is also a Brownian motion; (2) scaling: for all* $\lambda > 0$*, the process is also a Brownian motion; (3)Invariance by translation: for all* $s > 0$*, the process* $W_{t+s} - W_s)$ *is a standard Brownian motion independent of* $\mathcal{F}_s$*.*

In statistics, the expression $\sim (\mu, \Sigma)$ abbreviates a distribution with mean vector μ and covariance matrix Σ. $N(\mu, \Sigma)$ denotes a multivariate normal distribution with mean vector μ and covariance matrix Σ. Convergence in distribution and is asymptotically distributed are denoted as d and *plim* abbreviates the probability limit. The random variable independent and identically distributed is abbreviated is in short *i.i.d.*. A stochastic process u_t is called *white noise* if the u_t are i.i.d. with mean zero, $E(u_t) = 0$, positive definite covariance matrix $\Sigma = E(u_t, u_t)$, and finite fourth-order moments.

Definition 2.3. **?** *A process X is called adapted to the filtration* F(F$_t$)*, if for all t,* $X(t)$ *is F - measurable.*

The literature of the theory of stochastic differential equations is very large Kushner and P.Dupuis (2001),Oksendal (2000),**?**,**?**,

Malliavin and Thalmaien (2006),**?**, **?**,Gawarecki and Mandrekar (2011). Let $X(t)$ be a diffusion in n dimensions described by the multi - dimensional stochastic differential equation

$$dX(t) = \Phi(X(t), t)dt + \Psi(X(t), t)dB(t), \quad X(0) = x_0, \tag{1}$$

where Ψ is $n \times d$ matrix valued function, B is d-dimensional Brownian motion and and X and Φ are vector n- dimensional vector valued functions. The vector $\Phi(X, t)$ and the matrix $\Psi(X, t)$ are the coefficients of the stochastic differential equation.

Theorem 2.1. *(Unique and Existence of Solution) If the coefficients are locally Lipschitz in X with a constant independent of t, that is, for every N, there is a constant K depending only on T and N such that for all* $\mid x \mid, \mid y \mid \leq N$ *and all* $0 \leq t \leq T$*,*

$$\mid \Phi(x, t) - \Phi(y, t) \mid + \mid \Psi(x, t) - \Psi(y, t) \mid \leq K \mid x - y \mid, \tag{2}$$

for for any given $X(0)$ *the strong solution to stochastic differential equation 1 is unique. If in addition to condition 2 the linear growth condition holds*

$$\mid \Phi(x, t) \mid + \mid \Psi(x, t) \mid \leq K_\tau(1 + \mid x \mid), \tag{3}$$

$X(0)$ *is independent of B, and* $E \mid X(0) \mid^2 < \infty$*, then the strong solution exists and is unique on* $[0, T]$*, moreover,*

$$E(sup \mid X(t) \mid^2) < C(1 + E \mid X(0) \mid^2),$$

where constant C depends only on K and T.

The following theorem gives the solution of stochastic differential equations as Markov processes.

Theorem 2.2. **?** *(The solution of SDEs as Markov processes) If equation 1 satisfies the conditions of the existence and uniqueness theorem 2.1, the solution X_t of the equation for arbitrary initial values is a Markov process on the interval $[t_0, T]$ whose initial probability distribution at the instant to is the distribution of C and whose transition probabilities are given by*

$$P(s, x_t, B) = P(X_t \in B \mid X_s = x) = P(X_t(s, x) \in B] \tag{4}$$

where $X_t(s, x)$ is the solution of equation.

Theorem 2.3. **?** *Let $X(t)$ be a regular adapted such that with probability one $\int_0^T X^2(t)dt < \infty$. Then Ito integral $\int_0^T X(t)dB(t)$ is defined and has the following properties.*

1. *Linearity. If Itô integrals of $X(t)$ and $Y(t)$ are defined and α and β are some constants then $\int_0^T (\alpha X(t)+\beta Y(t))dB(t) = \alpha \int_0^T X(t)dB(t) + \beta \int_0^T Y(t)dB(t)$*
2. *$\int_0^T X(t)I_{(a,b]}(t)dB(t) = \int_a^b X(t)dB(t)$. The following two properties hold when the process satisfies an additional assumption*

$$\int_0^T E(X^2(t))dt < \infty. \tag{5}$$

3. *Zero mean property. If condition (5) holds then $E\Big(\int_0^T X(t)dB(t)\Big) = 0$.*
4. *Isometry property. If condition (5) holds. Then*

$$E\Big(\int_0^T X(t)dB(t)\Big)^2 = \int_0^T E\Big(X^2(t)\Big)dB(t)$$

Corollaire 2.3.1. *If X is a continuous adapted process then the Itô integral $\int_0^T X(t)dB(t)$ exists. In particular, $\int_0^T f(B(t))dB(t)$ where f is a continuous function on **R** is well defined.*

A consequence of the isometry property is the expectation of the product of two Itô integrals.

Theorem 2.4. *Let $X(t)$ and $Y(t)$ be regular adapted processes, such that $\int_0^T X(t)^2 dt < \infty$ and $\int_0^T Y(t)^2 < \infty$. Then,*

$$E\Big(\int_0^T X(t)dB(t) \int_0^T Y(t)dB(t)\Big) = \int_0^T E\Big(X(t)Y(t)\Big)dt. \tag{6}$$

We denote by $\mathbb{R}^{mn}$ all real - valued $m \times n$ matrices and by $W(t) = (W_1(t), ..., W_n(t))^{'}, t \geq 0$. Let $[a, b] \in [0, \infty[$ and we put $C_W([a, b]) = \{f : [a, b] \times \Omega \to \mathbb{R}^{mn}|$
$\forall\ 1 \leq i \leq m, \forall 1 \leq j \leq n : f_{ij} \in C_{Wj}([a, b])\}$ and $C_{IW}([a, b]) = \{f : [a, b] \times \Omega \to \mathbb{R}^{mn}|$
$\forall\ 1 \leq i \leq m, \forall 1 \leq j \leq n : f_{ij} \in C_{IWj}([a, b])\}$ and $C_I([a, b])$ respectively. If $f : [a, b] \times \Omega \to \mathbb{R}^{mn}$ belongs to $C_{IW}([a, b])$, then the stochastic integral with respect to W is the m - dimensional vector defined by $\int_a^b f(t)dW(t) = (\sum_{j=1}^n \int_a^b f_{ij}(t)dW_j(t))^{'}_{1\leq i\leq m}$ where each of the integrals on the right - hand side is defined in the sense of Itô.

Definition 2.4. **?** *If $f : [a, b] \times \Omega \to \mathbb{R}^{mn}$ belongs to $C_{IW}([a, b])$, then the stochastic integral with respect to W is the m - dimensional vector defined by*

$$\int_a^b f(t)dW(t) = (\sum_{j=1}^n \int_a^b f_{ij}(t)dW_j(t))^{'}_{1\leq i\leq m} \tag{7}$$

where each of the integrals on the right - hand side is defined in the sense of Itô.

Definition 2.5. *(Multiple Itô integral) Assuming that $W(t)$ is a standard Brownian motion,*

$$\frac{1}{n!}\int_0^t \int_0^{t_n} ... \int_0^{t_2} dW(t_1)...dW(t_n) = t^{n/2} H_n\Big(\frac{W(t)}{\sqrt{t}}\Big) \tag{8}$$

where $H_n(x) = (-1)^n e^{x^2/2} \frac{d^n}{dx^n} e^{-x^2/2}$ and $n! = n \times (n-1) \times (n-2) \times ... \times 1$. Here H_n is the Hermite polynomial.

It is important to get in mind that

$$\int_0^T W(t)dW(t) = \frac{1}{2}(W^2(T) - T), \tag{9}$$

where $dW(t_j) = \triangle W_j = W(t_{j+1}) - W(t_j)$ and $dt = \triangle t_j = t_{j+1} - t_j$.

3. Multidimensional Ornstein - Uhlenbeck Process

In mathematics, the Ornstein - Uhlenbeck process is a stochastic process with applications in the physical sciences and financial mathematics. Its original application in physics was as a model for the velocity of a massive Brownian particle under the influence of friction Uhlenbeck and Ornstein (1930), Wang and Uhlenbeck (1945). It is named after a Dutch physicist Leonard Salomon Ornstein (1880 - 1941) and a Dutch - American theoretical physicist George Eugene Uhlenbeck (1900 - 1988).

Also, the Ornstein-Uhlenbeck process is a stationary Gauss - Markov process, which means that it is a Gaussian process, a Markov process, and is temporally homogeneous. In fact, it is the only nontrivial process that satisfies these three conditions, up to allowing linear transformations of the space and time variables. Over time, the process tends to drift towards its mean function: such a process is called mean-reverting.

3.1 Model Specification

Let define the multivariate Ornstein - Uhlenbeck process as follows As mentioned in **?**,

$$dX(t) = -AX(t)dt + BdW(t), \tag{10}$$

where

$$X(t) = \begin{bmatrix} X_1(t) \\ ... \\ X_n(t) \end{bmatrix}, W(t) = \begin{pmatrix} W_1(t) \\ ... \\ W_m(t) \end{pmatrix},$$

$$A = \begin{pmatrix} a_{11} & ... & a_{1m} \\ ... & ... & ... \\ a_{n1} & ... & a_{nm} \end{pmatrix}, B = \begin{pmatrix} b_{11} & ... & b_{1m} \\ ... & ... & ... \\ b_{n1} & ... & b_{nm} \end{pmatrix}.$$

The matrices A and B are called the matrix of diffusion and volatility respectively and supposed to be positive definite matrices.

Theorem 3.1. *Let $dX(t) = -AX(t)dt + BdW(t)$ be an n-dimensional Itô process as above. Let $g(t, x) = (g_1(t, x), ..., g_p(t, x))$ be a C^2 map from $[0, \infty) \times \mathbb{R}^n$ into $\mathbb{R}^p$. Then the process $Y(t, w) = g(t, X(t))$ is again an Itô process, whose component number k, Y_k, is given by*

$$dY_k = \frac{\partial g_k}{\partial t}(t, X)dt + \sum_i \frac{\partial g_k}{\partial x_i}(t, X)dX_i + \frac{1}{2}\sum \frac{\partial^2 g_k}{\partial x_i \partial x_j}(t, X)dX_i dX_j, \tag{11}$$

where $dB_i dB_j = \delta_{ij}dt$ and $dB_i dt = dtdB_i = 0$.

3.2 Assumptions

Suppose that the R^n- valued function f and the $(d\times m)$ matrix - valued function G are assumed to be defined and measurable on $[t_0, T] \times R^n$. and have the following properties: there exists a constant $K > 0$ such that

A1. (*Lipschitz condition*).For all $t \in [t_0, T]$, $x \in R^d$,
$|f(t, x) - f(t, y)| + |G(t, x) - G(t, y)| \leqq K|x - y|$
The Lipschitz condition guarantees that $f(t, x)$ and $G(t, x)$ do not change faster with change in x than does the function X itself. This implies in particular the continuity of $f(t, x)$ and $G(t, x)$ for all $t \in [t_0, T]$.

A2. (*Restriction on growth*). For all $t \in [t_0, T]$, $x \in R^d$
$|f(t, x)|^2 - |G(t, x)|^2 \leqq K^2(1 + |x|^2)$.

A3. Suppose that the process $X = \{X_t\}_{t\geq 0}$ satisfies the following condition. For all $T > 0$ there exist positive constants α, β, D such that

$$E[| X_t - X_s |^\alpha \leq D \mid t - s |^{1+\beta}; 0 \leq s, t \leq T.$$

Then there exists a continuous version of X.

Then, equation (16) has on $[t_0, T]$ a unique R^d-valued solution X_t, continuous with probability 1, that satisfies the initial condition $X_{t_0} = C$; that is, if X_t and Y_t are continuous solution of the equation (16) with the same initial value C, then

$$P[sup|X_t - Y_t| > 0] = 0, \quad as \quad t_0 \leqq t \leqq T < \infty.$$

3.3 Multiple Itô Integral

Definition 3.1. *Let $B_t(w)$ be n-dimensional Brownian motion. Then we define $\mathcal{F}_t = \mathcal{F}_t^{(n)}$ to be the σ - algebra generalized by the random variables $B_s(.)$; $s \leq t$. In other words, $\mathcal{F}_t$ is the smallest σ - algebra containing all sets of the form*

$$\{w; B_t(w) \in \mathcal{F}_1, ..., B_{t_k}(w) \in F_k\},$$

where $t_j \leq t$ and $F_j \subset \mathbb{R}^n$ are Borel sets, $j \leq k = 1, 2, ...$ (We assume that all sets of measure zero are included in $\mathcal{F}_t$).

Let $\{(W(t), F_t)\}_{t\in[0,T]}$ be an m - dimensional Brownian motion with components $W_i(t), i = 1, ..., m$. Let $\{(X(t), F_t)\}_{t\in[0,T]}$ be an $\mathbb{R}^{n,m}$ - valued progressively measurable process with each component $X_{ij} \in \mathbb{L}^2[0, T]$. Then the multidimensional Itô integral of X with respect to W is defined in KornKornKroisandt2010 by

$$\int_0^t X(s)dW(s) := \begin{bmatrix} \sum_{j=0}^m \int_0^t X_{1j}(s)dW_j(s) \\ ... \\ \sum_{j=0}^m \int_0^t X_{nj}(s)dW_j(s) \end{bmatrix}, t \in [0, T], \tag{12}$$

where all single integrals inside the sums of the right - hand side are one - dimensional Itô integrals.

Let X and Y be two real - valued Itô processes with representations
$X(t) = X(0) + \int_0^t G(s)ds + \int_0^t H(s)dW(s)$ and $Y(t) = Y(0) + \int_0^t N(s)ds + \int_0^t M(s)dW(s)$, then, the quadratic covariation of X and Y is defined by

$$\langle X, Y\rangle_t := \sum_{i=0}^m \int_0^t H_i(s)M_i(s)d(s). \tag{13}$$

Theorem 3.2. *(Multidimensional Itô Formula) Let $X(t) = (X_1(t), X_2(t), ..., X_n(t))$ be an n-dimensional Itô process with*

$$X_i(t) = X_i(0) + \int_0^t K_i(s)ds + \sum_{j=1}^m \int_0^t H_{ij}(s)dW_j(s), \ i = 1, ..., n, \tag{14}$$

where $W(t)$ an m-dimensional Brownian motion. Let further $f : [0, \infty) \times \mathbb{R}^n \to \mathbb{R}$ be a $C^{1,2}$–function, i.e. f is continuous, continuously differentiable with respect to the first variable (time) and twice continuously differentiable with respect to the last n variables(space). We then have

$$\begin{aligned} & f(t, X_1(t), X_2(t), ..., X_n(t)) = f(0, X_1(0), X_2(0), ..., X_n(0)) \\ & + \int_0^t f_t(s, X_1(s), X_2(s), ..., X_n(s))ds + K_i(s)ds \\ & + \sum_0^n \int_0^t f_{x_i}(s, X_1(s), X_2(s), ..., X_n(s))dX_i(s) \\ & + \frac{1}{2} \sum_{i,j=0}^n \int_0^t f_{x_i} x_j(s, X_1(s), X_2(s), ..., X_n(s))d\langle X_i, X_j\rangle_s. \end{aligned} \tag{15}$$

3.4 From Multidimensional Ornstein - Uhlenbeck Process to VAR Process

In the economics analysis, the results of observations are usually interpreted as values of random variables. Thus the total data or the overall observation X is an n-dimensional random vector. In general, an observation is a random variable X with values in a measurable space. This means that there exists a probability space $(\Omega, \mathcal{F}, \mathbf{P})$. where n is the set of elementary outcomes;$\mathcal{F}$ is a $\sigma - algebra$ of events and $\mathbf{P}$ is the probability measure) and X is a measurable mapping of $\{\Omega, \mathcal{F}\} \to \{\mathcal{X}, \mathcal{U}\}$. The basic characteristic of a random variable X is its distribution which is the measure $\mathbf{P}^X$, given by $\mathbf{P}^X\{A\} = \mathbf{P}\{X \in A\}$ defined on $\mathcal{U}$.

$$dX(t) = -AX(t)dt + BdW(t). \tag{16}$$

As mentioned above, A and B are $n \times n$ constant matrices, positive definite matrices, and symmetric matrices with all positive eigenvalues. W denotes the vector of Brownian process for which the solution is

$$X(t) = X(0)exp(-At) + B\int_0^t exp[-A(t-s)]dW(s), \tag{17}$$

where exp denotes exponential symbol. By using backward - looking analysis, we write this model as follows:

$$X_t = X_{t-1}exp(-A) + B\int_0^t exp[-A(t-s)]dW(s). \tag{18}$$

Here we say that the coefficient matrix of $n \times n$, $\Pi = exp(-A)$, can be a real or complex constant matrix and positive definite matrix and the stochastic part of this model, $U_t = B\int_0^t exp[-A(t-s)]dW(s)$. Thus, the VAR(1) model derived in our analysis is given as follows

$$X_t = \Pi X_{t-1} + U_t, \quad U_t \sim N(0_{n\times n}, \Omega_{n\times n}). \tag{19}$$

In the time series analysis, the vector U_t is the vector of shocks so called innovations Lutkepohl (2005),**?**, Hamilton (1994). In our analysis, we will estimate the coefficients of the model (19).

4. Bayesian Vector Autoregressive Model

In time series analysis, the Vector Autoregressive models attract interest of many researchers in many fields such as Economics, Econometrics, Finance, Geoscience, Physics, Biology, etc. Berger (1985), Zellner (1971), Koop and Korobilis (2013),Koop and Korobilis (2013). The Bayesian vector autoregressive models are standard multivariate autoregressive models estimated by Bayesian estimation methods and are used in empirical investigations – structural analysis, and scenario analysis – and forecasting Litterman (1986),Sims (1980). The most commonly used multivariate time series model is the vector autoregressive model, particularly so in the econometric literature for good reasons. First, the estimation of model is very easy. Second, the literature the properties of BVAR models have been studied and developed. Finally, BVAR models are similar to the multivariate multiple linear regressions widely used in multivariate statistical analysis.

For example, in Economics, the pioneering work of Sims (1980) proposed to replace the large scale macroeconomic models popular in the 1960s with VARs, and suggested that Bayesian methods could have improved upon frequentist ones in estimating the model coefficients. Bayesian methods are increasingly becoming attractive to researchers in many fields such as Econometrics Koop2003. Bayesian VARs (BVARs) with macroeconomic variables were first employed in forecasting by Litterman (1986) but now it is one of the most policy analysis tools used by scholars and policy makers such as research institutions, central banks,and governments.

Suppose X_t is a zero mean, stationary Gaussian VAR(1) process of the form

$$X_t = \Pi X_{t-1} + U_t, \quad U_t \sim N(0_{n\times n}, \Omega_{n\times n}), \tag{20}$$

where U_t is a vector of innovations and the prior distribution for $\Theta := vec(\Pi)$ is a multivariate normal with known mean Θ^* and covariance matrix Ω_Θ. For the reasons of simplicity and practice, stationary, stable VAR(1) process has been considered. As well known in time series literature, a process is stationary if it has time invariant first and second moments. Since X_t follows a VAR(1) model, the condition for its stationariness is that all solutions of the determinant equation $\mid I_{kp} - \Phi B \mid = 0$ must be greater that 1 in modulus or they are outside the unit circle Lutkepohl (2005), **?**, Hamilton (1994), citeSims1980.

The multivariate time series X_t follows a vector autoregressive model of order p, VAR(p), that is a generalization of a vector autoregressive model of order 1, VAR(1), if

$$X_t = \Phi_0 + \sum_{i=1}^{p} \Phi_i X_{t-i} + \epsilon_t, \quad \epsilon_t \sim N(0_{n\times n}, \Sigma_{n\times n}), \tag{21}$$

where Φ_0 is a k-dimensional constant vector and Φ_i are $k \times k$ matrices for $i > 0$; $\Phi_p \neq 0$, and ϵ_t is a sequence of independent and identically distributed ($i.i.d.$) random vectors with mean zero and covariance matrix Σ_ϵ, which is positive

- definite. Also,the necessary and sufficient condition for the weak stationary of the VAR(p) series is that all solutions of the determinant equation $\mid \Psi(B) \mid = 0$ must be greater than 1 in modulus.

We can express a VAR(p) model in a stationary VAR(1) form by using an expanded series.

Define $Y_t = (X_t^{'}, X_{t-1}^{'}, ..., X_{t-p+1}^{'})^{'}$, which is pk- dimensional time series. The VAR(p) in Equation 19 can be written as

$$Y_t = \Xi Y_{t-1} + \Lambda_t, \tag{22}$$

where $\Lambda_t = (U_t^{'}, 0^{'})$ with 0 being a $k(p-1)$ - dimensional zero vector, and

$$\Xi = \begin{bmatrix} \psi_1 & \psi_2 & ... & \psi_{p-1} & \psi_p \\ I & 0 & ... & 0 & 0 \\ 0 & I & ... & 0 & 0 \\ .... & ... & ... & ... & ... \\ 0 & 0 & ... & I & 0 \end{bmatrix}, \tag{23}$$

where it is understood that I and 0 are the $k \times k$ identity and zero matrix, respectively. The matrix Ξ is called the companion matrix of the matrix polynomial $\Psi(B) = I_k - \Psi_1 B_1 - ... - \Psi_p B_p$. The covariance matrix of Λ_t has a special structure; all of its elements are zero except those in the upper-left corner that is Σ_ϵ.

In multivariate time series analysis and econometric analysis, the useful tools used by policy-makers are the forecast error variance decomposition and the impulse response functions.

The Moving Average with the infinity lags ($MA(\infty)$) representation of the VAR(p) model is given by

$$X_t = \mu + \sum_{i=1}^{\infty} \Psi_i \epsilon_{t-i}, \tag{24}$$

where Ψ_i denotes the coefficients matrix that measures the impact of innovations on X_t at time t and ϵ_t is a vector of innovations so-called in literature 'shocks'. In economic analysis, we have many shocks that influence the economy's fluctuations like demand shock, supply shock, monetary policy shock, fiscal policy shock, energy shock, war shock, political instability shock, population shock, health shock, environment shock and so on.

4.1 Bayesian Estimation Methods for VAR Models

In the Bayesian approach, as mentionned in Litterman (1986), Lutkepohl and Kratzig (2004), Lutkepohl (2005),Koop and Korobilis (2013), let us assume that (i) the non sample or prior information is available in the form of a density. Denoting the parameters of interest by Θ, (ii) the prior information is given in the prior probability density function (p.d.f.) $g(\Theta)$; (iii) the sample information is in the sample, say $f(y \mid \Theta)$, which is algebraically identical to the likelihood function $\mathcal{L}(\Theta \mid X)$, The two types of information are combined in the Bayes' theorem states

$$g(\Theta \mid X) = \frac{f(X \mid \Theta) g(\Theta)}{f(X)} \tag{25}$$

where $f(X)$ denotes the unconditional sample density which, for a given sample, is just a normalizing constant. In other words the distribution of Θ, given the sample information contained in X, can be summarized by $g(\Theta \mid X)$. This function is proportional to the likelihood function times the prior density $g(\Theta)$,

$$g(\Theta \mid X) \propto f(X \mid \Theta) g(\Theta) = \mathcal{L}(\Theta \mid X) g(\Theta) \tag{26}$$

The conditional density $g(\Theta \mid X)$ is the posterior p.d.f.. It contains all the information available on the parameter vector Θ. Point estimators of Θ may be derived from the posterior distribution. That is,

$$posterior \ \ distribution \propto likelihood \ \ \times prior \ \ distribution. \tag{27}$$

The normal prior for the parameters of a Gaussian VAR model, $\Theta := vec(A) = vec(A_1, ..., A_p)$ is a multivariate normal with known mean Θ^* and covariance matrix Ω_θ,

$$g(\Theta) = \Big(\frac{1}{2\pi}\Big)^{K^2 p/2} \mid \Omega_\theta \mid^{-1/2} exp\Big[-\frac{1}{2} (\Theta - \Theta^*)^{'} \Omega_\theta^{-1} (\Theta - \Theta^*) \Big]. \tag{28}$$

The Gaussian likelihood function

$$g(\Theta) = \Big(\frac{1}{2\pi}\Big)^{KT/2} \mid I_T \otimes \Sigma_u \mid^{-1/2}$$
$$\times\, exp\Big[-\frac{1}{2}(X - (Z^{'} \otimes I_T)\Theta)^{'}(I_T \otimes \Sigma_u^{-1})(X - (Z^{'} \otimes I_T)\Theta)\Big]. \tag{29}$$

Combining the prior information with the sample information summarized in the Gaussian likelihood function gives the posterior density

$$g(\Theta) \propto \mathcal{L}(\Theta \mid X)g(\Theta)$$
$$\propto exp\Big\{ -\frac{1}{2}[(\Omega_\theta^{-1/2}(\Theta-\Theta^*)^{'}(\Omega_\theta^{-1/2}(\Theta-\Theta^*)]$$
$$+\{(I_T \otimes \Sigma_u^{-1})X - (Z^{'} \otimes \Sigma_u^{-1/2})\Theta\}^{'}\{(I_T \otimes \Sigma_u^{-1})X - (Z^{'} \otimes \Sigma_u^{-1/2})\Theta\}]\Big\}. \tag{30}$$

Here $\Omega_\theta^{-1/2}$ and $\Sigma_u^{-1/2}$ denote the symmetric square root matrices of Ω_θ^{-1} and Σ_u^{-1}, respectively. The white noise covariance matrix Σ_u is assumed to be known for the moment. Defining $w^{'} := [\Omega_\theta^{-1/2}\Theta^* \quad (I_T \otimes \Sigma_u^{-1})X]^{'}$ and $W^{'} := [\Omega_\theta^{-1/2} \quad Z^{'} \otimes \Sigma_u^{-1}]^{'}$, the exponent in (30) can be rewritten as

$$-\frac{1}{2}(w - W\Theta)^{'}(w - W\Theta)$$
$$= -\frac{1}{2}\Big[(\Theta - \bar{\Theta})^{'}W^{'}W(\Theta - \bar{\Theta}) + (w - W\bar{\Theta})^{'}(w - W\bar{\Theta})\Big],$$

where

$$\bar{\Theta} := (W^{'}W)^{'}Ww = [\Omega_\theta^{-1} + (ZZ^{'} \otimes \Sigma_u^{-1})]^{-1}[\Omega_\theta^{-1}\Theta^* + (Z^{'} \otimes \Sigma_u^{-1})X]$$

The final values of the parameters obtained in the computation are called the posterior mean of VAR(1) coefficients estimated by using Minnesota.

4.1.1 Asymptotic Properties of Bayesian Estimators

We describe the asymptotic properties of the Bayesian estimators with the help of another general result.

Theorem 4.1. *Let $\hat{\vartheta}_T$ be a family of Bayesian estimators and prior density $p(.) \in \mathcal{P}_c$. Assume that the normalized likelihood ratio $Z_{T,\vartheta}(.)$ possess the following properties:*

1. *For any compact $\mathbb{K} \subset \theta$ there correspond numbers $\alpha(\mathbb{K}) = \alpha$ and $\beta(\mathbb{K}) = \beta$ and $\kappa_T^{\mathbb{K}} = \kappa_T(.) \subset \Phi$, such that For some $q > 0$, all $|u_1| < R$,$|u_2| < R$ and any $R > 0$ $\sup \mathbf{E}_\theta |Z_{T,\theta}(u_2)^{\frac{1}{2}} - Z_{T,\theta}(u_1)^{\frac{1}{2}}|^2 \le \beta(1 + R^\alpha)|_2 - u_1|^q$ for all $u \in U_{T,\theta}$ $\sup \mathbf{E}_\theta |Z_{T,\theta}(u)^{\frac{1}{2}} \le exp(-k_T(|u|)$.*
2. *The marginal distributions of the random $Z_{T,\theta}(u)$ uniformly in $\theta \in \mathbb{K}$ convergence to marginal distributions of the random functions $Z_\theta(u)$.*
3. *$\psi(\upsilon) = \int \mathcal{L}(\upsilon - u)\frac{Z_\theta(u)}{\int_{\mathbb{R}}^d Z_\theta(y)dy}du$*

with probability 1 attains its absolute minimum value at the unique point $\bar{u}(\theta) = \bar{u}$. Then the Bayesian estimator $\tilde{\theta}_T$ is consistent uniformly in $\theta \in \mathbb{K}$,i;e;, for any $\upsilon > 0$ $\lim\ supP_\theta^{(T)}\Big\{|\tilde{\theta}_T - \theta| > \upsilon\Big\} = 0$, the distribution of the random variables $\tilde{u} = \varphi_T(\theta)^{-1}(\tilde{\theta}_T - \theta)$ converge uniformly in $\theta \in \mathbb{K}$ to the distribution of $\tilde{u}(\theta) = \tilde{u}$ and for any loss function $\mathcal{L}(.) \in W_p$ we have uniformly in $\theta \in \mathbb{K}$ $\lim \mathbf{E}_\vartheta \mathcal{L}(\varphi_T(\theta)^{-1}(\tilde{\vartheta}_T - \vartheta)) = \mathbf{E}_\vartheta \mathcal{L}(\tilde{u})$.

Theorem 4.2. *Under some conditions. The Bayesian estimator $\hat{\vartheta}_T$ is uniformly consistent on compacts $\mathbb{K} \subset \Theta$,i.e., for any $v > 0$, lim sup $\mathbb{P}_\vartheta^{(T)}\Big\{ \mid \hat{\vartheta}_T - \vartheta \mid> v\Big\} = 0$, uniformly asymptotically normal, $\mathcal{L}_\vartheta\Big\{T^{\frac{1}{2}}(\hat{\vartheta}_T - \vartheta)\Big\} \Rightarrow N(0, I(\vartheta)^{-1})$, and the moment converge: for any $p > 0$ uniformly on compact $\mathbb{K}$ $\lim \boldsymbol{E}_\vartheta \mid T^{1/2}\hat{\vartheta}_T - \vartheta) \mid^p = \boldsymbol{E} \mid I(\vartheta)^{-1/2}\xi \mid^p$, $\xi \sim N(0, J)$, as $T \to \infty$, where J is a unit $d \times d$ matrix. Moreover the Bayesian estimator is asymptotically efficient for loss functions $l \in W_p$.*

4.1.2 Selection of Bayesian VAR Model

Before estimating VAR(p) model, see par example some works Koop (2003) Sims (1980), it is important to determine the optimal lag of this model. In practice, we use some information criteria like the Akaike information Criterion (AIC), the Schwarz information Criterion (SC), and the Hannan - Quinn information Criterion (HQ). These criteria may be used in this formula $AIC = ln\ \tilde{\sigma}^2 + \frac{2}{T}(k)$, $HQ = ln\ \tilde{\sigma}^2 + \frac{2\ ln\ lnT}{T}(k)$, $SC = ln\ \tilde{\sigma}^2 + \frac{ln\ T}{T}(k)$. Here $\tilde{\sigma}^2$ stands for the sum of squared estimation residuals divided by the sample size T and k is a number of estimated parameters.

4.2 Structural Analysis

For policy-making decisions, we use the forecast error variance decomposition and the impulse response functions to analyze the impact of the innovations of one variable on another variable.

4.2.1 Forecast Error Variance Decomposition

Using the MA representation of a VAR(p) model and the fact that $Cov(\eta_t) = I_k$, we see that the $l - step$ ahead error of Z_{h+l} at the forecast origin $t = h$ can be written as

$$e_h(l) = \psi_0\eta_{h+l} + \psi_1\eta_{h+l-1} + ... + \psi_{l-1}\eta_{h+l}, \tag{31}$$

and the covariance matrix of the forecast error is

$$Cov[e_h(l)] = \sum_{v=0}^{l-1} \psi_v\psi_v^{'}. \tag{32}$$

From Equation (32), the variance of the forecast error $e_{h,i}(l)$, which is the ith component of $e_h(l)$ is

$$Var[e_{h,i}(l)] = \sum_{v=0}^{l-1}\sum_{j=1}^{k} \psi_{v,ij}^2 = \sum_{j=1}^{k}\sum_{v=0}^{l-1} \psi_{v,ij}^2. \tag{33}$$

Using Equation (), we define

$$\omega_{ij}(l) = \sum_{v=0}^{l-1} \psi_{v,ij}^2, \tag{34}$$

and obtain

$$Var[e_{h,i}(l)] = \sum_{j=1}^{k} \omega_{ij}(l). \tag{35}$$

Therefore, the quantity $w_{ij}(l)$ can be interpreted as the contribution of the jth shock η_{jt} to the variance of the $l - step$ ahead forecast error of Z_{it}. Equation () is referred to as the forecast error decomposition. In particular, $\omega_{ij}(l)/Var[e_{h,i}(l)]$ is the percentage of contribution from the shock η_{jt}.

4.2.2 Impulse Response Functions

In multivariate time series analysis like in Lutkepohl and Kratzig (2004), Hamilton (1994),Tsay (2014) , we use the impulse response function when we need to know the impact of one variable on another. For example, one might be interested in knowing the effect on the monthly inflation rate if the monthly exchange rate growth will increase or decrease. In this analysis, the impulse response functions are used in the statistical and econometric literature. The coefficient matrix Ψ_i of the MA representation of a VAR(p) model is referred to as the coefficients of impulse response functions. The summation $\Phi_n = \sum_{i=0}^{n} \Psi_i$ denotes the accumulated responses over n periods to a unit shock to Z_t. From the MA representation of Z_t and using the Cholesky decomposition of Σ_ε, we have

$$Z_t = [\Phi_0 + \Phi_1 B + \Phi_2 B^2 + ...]\eta_t, \tag{36}$$

where $\Phi_l = \Psi_l U^{'}$, $\Sigma_\varepsilon = U^{'}U$, and $\eta_t = (U^{'})^{-1}\varepsilon_t$ for $l \geq 0$. Thus, components of η_t are uncorrelated and have unit variance. The total accumulated responses for all future periods are defined as $\Phi_\infty = \sum_{i=0}^{\infty} \Psi_i$ and called the total multipliers or long - run effects.

4.3 Forecasting

Forecasting is one of the most activities of policy-makers because it informs them how the policy decisions are and helps them to maintain a forward-focused mindset. In time series analysis, the VAR(p) models are used for this goal.

5. Empirical Analysis

In mathematical modeling, stochastic differential equations and vector autoregressive models are used in many fields such as Meteorology, Physics, Biology, Hydrology, Finance, and Economics. This section uses the Bayesian Vector Autoregressive model to analyze and forecast the three macroeconomic variables.

5.1 Economic Intuitions Behind the BVAR(1) Model

Macroeconometric modeling grew rapidly in importance during the late I950S and the I960s, going on to achieve a very influential role in macroeconomic policy-making during the I970s. Its failure to deliver the detailed economic control that it had seemed to promise then led to a barrage of attacks, ranging from disillusion and skepticism on the part of policymakers to detailed and well-argued academic criticism of the basic methodology of the approach.

Perhaps, the most powerful and influential of these academic arguments came from Sims (1980) in his article 'Macroeconomics and Reality'. Sims argued, on three quite separate grounds, against the basic process of model identification, which lies at the heart of the Cowles Commission methodology. Firstly that economic theory gives rise to identification restrictions which are typically more complex than those traditionally applied in macroeconometric models. In particular, he said that theory normally implies complex cross-equation restrictions that require system estimation and which cannot be imposed on a single equation basis. Secondly, traditional identification conditions are often met simply because of the presence of dynamics in the models. Sims (1980) argued that this identification is spurious and technically invalid, as purely dynamic terms cannot help in structural identification in the conventional sense. Finally, he said that the importance of expectations effects and the interaction of policy regimes and agents' expectations make identification very difficult Sims (1980) argued that any one of these problems would form a challenging, but feasible, research agenda, but that 'Doing all of these at once would be a program which is so challenging as to be impossible in the short run'. He then proposed a methodology based on vector autoregressive (VAR) models **?**. As argued by Christopher Sims, most economists would agree that there are many macroeconomic variables whose cyclical fluctuations are of interest, and would agree further that fluctuations in these series are interrelated Sims (1980).

5.2 Model Specification

To illustrate this approach, we assume that the stationary BVAR(1) process, time - invaryING parameters, and the economy is stayed in same cycle during the period of analysis takes the matrix form as follows

$$\begin{bmatrix} \pi_t \\ e_t \\ m_t \\ h_t \end{bmatrix} = \begin{bmatrix} \psi_{11} & \psi_{12} & \psi_{13} & \psi_{14} \\ \psi_{21} & \psi_{22} & \psi_{23} & \psi_{24} \\ \psi_{31} & \psi_{32} & \psi_{33} & \psi_{34} \\ \psi_{41} & \psi_{42} & \psi_{43} & \psi_{44} \end{bmatrix} \times \begin{bmatrix} \pi_{t-1} \\ e_{t-1} \\ m_{t-1} \\ h_{t-1} \end{bmatrix} + \begin{bmatrix} u_{\pi t} \\ u_{et} \\ u_{mt} \\ u_{ht} \end{bmatrix}. \quad (37)$$

where π_t, m_t,e_t, and h_t denote the monthly CPI inflation rate, the change of exchange rate, money growth, and the change of the cooper price, respectively.

5.3 Data Analysis

To illustrate Bayesian VAR(1) model using some of the informative priors such as Minnesota. We use monthly data from Democratic Republic of the Congo data set on inflation rate π_t, the change of exchange rate e_t, money growth m_t, and the change of the cooper price h_t. The sample runs from january 2004 to september 2018.

Table 1. Summary Statistics

	π_t	e_t	m_t	h_t
Mean	0.0	0.0	0.0	0.0
Median	0.0	-0.0	0.0	0.0
Max.	0.1	0.1	0.2	0.2
Min.	-0.1	-0.1	-0.1	-0.4
Std. Dev.	0.0	0.0	0.1	0.1
Skewness	1.6	0.6	0.2	-0.8
Kurtosis	9.4	6.9	3.1	7.7
Jarque - Bera Stat.	366.8	120.6	1.3	181.6
Prob.(JB)	0.0	0.0	0.5	0.0
Sum	2.2	1.5	3.4	0.9
Observations	176	176	176	176

5.4 Selection of BVAR(1) Model

By using the software Eviews 11 and the data, the maximum lag of estimated BVAR model is 1. Therefore, the structural analysis and forecasting will be done with the BVAR(1)process.

Table 2. Selection of optimal lag

Lag	log L	AIC	SC	HQ
0	1280.9	-15.2	-15.1	-15.2
1	1370.7	-16.1^*	-15.7^*	-15.9^*
2	1379.7	-16.0	-15.3	-15.7
3	1387.8	-15.9	-14.9	-15.5
4	1394.7	-15.8	-14.5	-15.3

(*) indicates the maximum calculated lag of the VAR(p) model.

5.5 Estimated Posterior Mean Coefficients of BVAR Model

Like other VAR models, the BVAR(1) has many estimated coefficients and is hard to interpret.

Table 3. Posterior mean of BVAR(1) model

	π_t	e_t	m_t	h_t
π_{t-1}	0.3	0.2	0.3	0.2
	[5.7]	[2.4]	[1.9]	[0.9]
e_{t-1}	0.3	0.3	0.1	0.2
	[5.8]	[4.3]	[0.5]	[1.4]
m_{t-1}	0.1	0.1	-0.1	0.0
	[3.0]	[3.1]	[-2.4]	[0.4]
h_{t-1}	0.0	-0.0	-0.0	0.3
	[1.0]	[-2.2]	[-0.9]	[4.9]

Table 3 presents posterior means of all the VAR coefficients for Minnesota and [...] denotes Student statistics used for testing the statistical significance of estimated parameters. Our results can be given in matrix form as mentioned above by

$$\hat{\Pi} = \begin{pmatrix} 0.3 & 0.3 & 0.1 & 0.0 \\ 0.2 & 0.3 & 0.1 & -0.0 \\ 0.3 & 0.1 & -0.1 & -0.0 \\ 0.2 & 0.2 & 0.0 & 0.3 \end{pmatrix}$$

5.6 Derivation of Drift and Stochastic Volatility Coefficients Matrices

For calculating drift matrix of the model,(**??**), we use the MATLAB R2016a Software to compute these matrices because it is difficult to do so manually. Recall $exp(-A) := exp(\hat{\Phi}) = \hat{\Pi}$, we calculate the logarithm of matrix $\hat{\Pi}$, $ln(\Pi)$. For matrix theory, see Gentle (2007), **?**,Magnus and Neudecker (2007),Turkington (2002),Graham (1981), Higham (2008),Lewis (2001),Seber (2008) ,**?**, Gantmacher (2000), we get

$$\Phi = ln(\Pi) \approx \sum_{k=1}^{\infty} (-1)^{k+1} \frac{(\Pi - I)^k}{k},$$

where I is an identity matrix, and by using this powerful algebraic computing Software

$$\hat{\Phi} = \begin{pmatrix} -1.2 + 0.0i & -1.3 + 0.0i & -2.79 + 0.0i & -4.09 + 0.0i \\ -1.8 + 0.0i & -1.4 + 0.0i & -2.5 + 0.0i & -3.10 + 3.1i \\ -1.2 + 0.0i & -2.7 + 0.0i & -2.0 + 3.1i & -3.15 + 3.1i \\ -1.8 + 0.0i & -1.5 + 0.0i & -4.0 + 0.0i & -1.28 + 0.0i \end{pmatrix}.$$

As mentioned above, the matrix of the drift coefficient can be gotten by $\hat{A} = -\hat{\Phi}$

$$\hat{A} = \begin{pmatrix} 1.2 + 0.0i & 1.3 + 0.0i & 2.8 + 0.0i & 4.1 + 0.0i \\ 1.8 + 0.0i & 1.4 + 0.0i & 2.5 + 0.0i & 3.1 - 3.1i \\ 1.3 + 0.0i & 2.7 + 0.0i & 2.0 - 3.1i & 3.2 - 3.1i \\ 1.8 + 0.0i & 1.5 + 0.0i & 4.0 + 0.0i & 1.3 + 0.0i \end{pmatrix}.$$

In statistics, the covariance matrix is a symmetric matrix, $a_{ij} = a_{ji}$. Therefore the estimated Bayesian VAR residual covariance matrix is also a symmetric matrix given by

$$\hat{\Omega}_u = \begin{pmatrix} 0.0003 & 0.0001 & 0.0000 & 0.0000 \\ 0.0001 & 0.0005 & 0.0001 & 0.0000 \\ 0.0000 & 0.0001 & 0.0035 & -0.0001 \\ 0.0000 & 0.0000 & 0.0001 & 0.0043 \end{pmatrix}.$$

We get the matrix of stochastic volatility or the diffusion coefficient $\hat{B}$ by solving the stochastic integral $\int_0^t exp[-A(t - t^{'})]BdW(t^{'})$, that is, $\hat{B} = \hat{\Pi}^{-1}\hat{\Omega}_u$

$$\hat{B} = \begin{pmatrix} 0.002 & -0.003 & -0.004 & -0.009 \\ -0.002 & 0.005 & 0.010 & 0.013 \\ 0.003 & -0.005 & -0.026 & -0.016 \\ 0.000 & -0.002 & -0.004 & 0.012 \end{pmatrix}.$$

As mentioned the matrix B is the matrix of coefficients diffusion called the matrix of stochastic volatility in (16)

5.7 Forecast Error Variance Decompositions

The variance decomposition of inflation shows that 82 % of its innovations are due to itself innovations and 13 %, 6 %, and 41 % are due to exchange rate, money, and cooper price index innovations. With 13 % the exchange rate contributes more the CPI inflation. The variance decomposition of exchange rate shows that 79 % of its innovations are due to itself innovations and 15 %, 3 %, and 2 % are due to inflation rate, money growth, and cooper price index innovations. The variance decomposition of money shows that 97 % of its innovations are due to itself innovations and 2 % , 1 % , and 27 points of percentage are due to inflation, exchange rate, and cooper price index innovations. The variance decomposition of inflation shows that 98 % of its innovations are due to itself innovations and 1 % , 1 % , and 16 points of percentage are due to inflation, exchange rate, and money innovations. For policy-makers, this results show that there is a close connexion between inflation rate and exchange rate because of dollarization of economy and extra-version of economy. This monetary phenomenon dated since 1990s where the Congolese economy fell down by the political and socio - economical crises and army conflicts.

5.8 Macroeconomic Forecasting

For policy - makers, the forecasting the economy perform the policy - making decision but it costs much and is risky, often humbling tasks. Unfortunately, they are the jobs that many statisticians, economists and others are required to engage in as mentioned in many works Litterman (1986), Koop and Korobilis (2013). Nowadays, in the most Central Banks and other institutions , VARs models have become powerful tools of forecasting. The outputs from Bayesian VAR models seem to be more accurate and robust than the outputs of other approaches.

Table 4. Macroeconomic Forecasts October 2018 - March 2019

Period	Inflation	Exchange rate	money growth	Cooper price
October 2018	0.01	0.01	0.02	0.01
November 2018	0.01	-0.00	0.02	0.01
December 2018	0.11	0.11	0.18	0.23
January 2019	-0.07	-0.10	-0.12	-0.35
February 2019	0.02	0.02	0.05	0.07
March 2019	1.55	0.60	0.20	-0.75

6. Discussion of Results

The main job of the mathematicians is to improve the policy-making process by providing the best forecasting and analysis to policy-makers. By combining the stochastic analysis, the time series analysis, and the Bayesian inference, this approach is useful to mathematical modelling. The stochastic differential equations put the uncertainty in the model and the Bayesian statistics puts it in the parameters. In the last six decades, the stochastic differential equations are more used in mathematical modelling that are applied in many fields of science. Since the works of Black and Scholes in 1973, in the Economics and Finance, the stochastic analysis is more used by the Econometricians for econometric analysis Black and Scholes (1973), Bergstrom and Nowman (2007). Therefore, the new field is born called Financial Econometrics. To move algebraical from the stochastic analysis to time series analysis is strong challenge. This approach asks the strong foundation in these fields. In the last four decades, the BVAR models are used to forecast the economical and financial data Kadiyala and Karlsson (1997), Bikker (1998), SljiviC (2017), Droumaguet et al. (2016), Kwon et al. (2008), Peters et al. (2010), Chan et al. (2020). Among other methods, the BVARs are considered as powerful tools of forecasting Carriero et al. (2012), Litterman (1986), Doan et al. (1984). The BVAR models are the 'a-theorical' models so - called 'Black-boxes' that because they do not have the theorical foundations.

7. Conclusion

This paper has shown the connexion between the multivariate stochastic differential equations and the Bayesian vector autoregressive models thoroughly the Itô integrals. These stochastic continuous - time models are more used and applied in financial econometrics because of their capacity to put together uncertainties that are inherent to real - world economic problems. Also, the Bayesian vector autoregressive models are more attractive because they estimate the posterior mean estimators that are the product of modelling knowledges and data behavioursSims (1980).

Our methodology, however, is more general and could be applied to the analysis of factors driving misspecification and uncertainties observed in real world. This double - counting of uncertainties - stochastic differential equations modelling and Bayesian vector autoregressive models - that figure out the uncertainty in the models and the parameters is more occurrence. Thus, this mix approach should not be considered useful in the modern econometrics but also in any applied scientific field.

For next future research, there are many topics where someone can work and get the best results such as the Ornstein - Uhlenbeck process with jump, the Ornstein - Uhlenbeck process with Markov switchning chains.

References

Amemiya, T. (1985). *Advanced Econometrics*. Harvard University Press, Cambridge.

Basawa, I. V., Heyde, C. C., & Taylor, R. L. (2001). *Seleted Proceedings of the symposium on Inference for Stochastic Processes*, volume 37, 2001. Institute of Mathematical Statistics.

Bergstrom, A. R., & Nowman, K. B. (2007). *A continuous time econometric model of the United Kingdom with stochastic trends.* Cambridge University Press.

Bikker, J. A. (1998). Inflation forecasting for aggregates of the EU-7 and EU-14 with Bayesian VAR models. *Journal of Forecasting, 17*(2), 147-165.

Black, F., & Scholes, M. (1973). The pricing of options and corporate liabilities. *Journal of political economy, 81*(3), 637-654.

Carriero, A., Kapetanios, G., & Marcellino, M. (2012). Forecasting government bond yields with large Bayesian vector autoregressions. *Journal of Banking & Finance, 36*(7), 2026-2047.

Chan, J. C., Eisenstat, E., & Strachan, R. W. (2020). Reducing the state space dimension in a large TVP-VAR. *Journal of Econometrics, 218*(1), 105-118.

Chin, K. H., & Li, X. (2019). Bayesian forecast combination in VAR-DSGE models. *Journal of Macroeconomics, 59*, 278-298.

David, W. L. (2001). *Matrix Theory*. World Scientific, London.

Doan, T., Litterman, R., & Sims, C. (1984). Forecasting and conditional projection using realistic prior distributions. *Econometric reviews, 3*(1), 1-100.

Droumaguet, M., Warne, A., & Woániak, T. (2017). Granger causality and regime inference in Markov switching VAR models with Bayesian methods. *Journal of Applied Econometrics, 32*(4), 802-818.

Gantmacher, F. R. (2000). The Theory of Matrices vol 1 (Providence, Rhode Island: AMS Chelsea).

Gawarecki, L., & Mandrekar, V. (2010). Stochastic differential equations in infinite dimensions: with applications to stochastic partial differential equations. Springer Science & Business Media.

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). Bayesian Data Analysis Chapman & Hall. *CRC Texts in Statistical Science.*

Gentle, J. E. (2007). Matrix algebra. *Springer texts in statistics.* Springer, New York, NY.

Geweke, J. (2005). *Contemporary Bayesian econometrics and statistics.* John Wiley & Sons.

Graham, A. (1981). *Kronecker Product and Matrix Calculus with Applications*. Ellis Horwood.

James, D. H. (1994). *Time Series Analysis*. Princeton University Press, Princeton.

James, O. B. (1985). Statistical decision theory and Bayesian analysis. *Springer Series in Statistics, ISBN-10: 0-387-96098-8 and-13, 978-0387.*

Jan, R. M., & Heinz, N. (2007). *Matrix Differential Calculus with Applications in Statistics and Econometrics*. John Wiley and Sons, 3 edition, 2007.

Kadiyala, K. R., & Karlsson, S. (1997). Numerical methods for estimation and inference in Bayesian VAR-models. *Journal of Applied Econometrics, 12*(2), 99-132.

Karatzas, I., & Shreve, S. E. (1988). Brownian Motion and Stochastic Calculus Springer Verlag, New York. MR917065.

Kelly, D., & Smith, C. (2011). Bayesian inference for probabilistic risk assessment: A practitioner's guidebook. Springer Science & Business Media.

Koop, G. (2003) *Bayesian Econometrics*. John Wiley and Sons, West Sussex.

Koop, G., & Korobilis, D. (2013). Large time-varying parameter VARs. *Journal of Econometrics, 177*(2), 185-198.

Kushner, H. J. K., Kushner, H. J., Dupuis, P. G., & Dupuis, P. (2001). Numerical methods for stochastic control problems in continuous time (Vol. 24). Springer Science & Business Media.

Kushner, H., & Yin, G. G. (2003). *Stochastic approximation and recursive algorithms and applications* (Vol. 35). Springer Science & Business Media.

Kwon, Y., Bozdogan, H., & Bensmail, H. (2008). Performance of model selection criteria in Bayesian threshold VAR (TVAR) models. *Econometric Reviews, 28*(1-3), 83-101.

Litterman, R. B. (1986). Forecasting with Bayesian vector autoregressionsłfive years of experience. *Journal of Business & Economic Statistics, 4*(1), 25-38.

Lutkepohl, H. (2005). *New Introduction to Multiple Time Series Analysis*. Springer - Verlag, Berlin and Heidelberg,

Ltkepohl, H., & Krätzig, M. (Eds.). (2004). Applied time series econometrics. Cambridge university press.

Malliavin, P., & Thalmaier, A. (2006). *Stochastic calculus of variations in mathematical finance* (pp. 3-15). Berlin: Springer.

Nicholas, J. H. (2008). *Functions of Matrices Theory and Compuations*. SIAM.

O'Hagan, A., & West, M. (Eds.). (2010). *The Oxford handbook of applied Bayesian analysis*. OUP Oxford.

O'Hagan, A., & West, M. (Eds.). (2010). *The Oxford handbook of applied Bayesian analysis*. OUP Oxford.

Oksendal, B. (2000). *Stochastic Differential Equations*. Springer-Verlag, Berlin, 5 edition.

Peters, G. W., Kannan, B., Lasscock, B., & Mellen, C. (2010). Model selection and adaptive Markov chain Monte Carlo for Bayesian cointegrated VAR model. *Bayesian Analysis, 5*(3), 465-492.

Rao, B. P. (2011). Statistical inference for fractional diffusion processes.

Robert, C. P. (2007). *The Bayesian choice: from decision-theoretic foundations to computational implementation* (Vol. 2). New York: Springer.

Seber, G. A. (2008). *A matrix handbook for statisticians*. John Wiley & Sons.

Sim, C. H. (1990). First-order autoregressive models for gamma and exponential processes. *Journal of Applied Probability, 27*(2), 325-332.

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica: journal of the Econometric Society*, 1-48.

Šljivić, N. M. (2017). Cross-entropy method for estimation of posterior expectation in Bayesian VAR models. *Communications in Statistics-Theory and Methods, 46*(23), 11933-11947.

Ruey, S. T. (2014). *Multivariate Time Series Analysis*. Wiley, Hoboken: NJ.

Turkington, D. A. (2005). *Matrix calculus and zero-one matrices.* Cambridge Books.

Uhlenbeck, G. E., & Ornstein, L. S. (1930). On the theory of the Brownian motion. *Physical review, 36*(5), 823.

Wang, M. C., & Uhlenbeck, G. E. (1945). On the theory of the Brownian motion II. *Reviews of modern physics, 17*(2-3), 323.

Yury, A. K. (1998). *Statistical Inference for Spatial Poisson Processes*. Springer Science+Busmess Media, New York.

Yury, A. K. (2004). *Statistical Inference for ergodic Diffusion Processes*. Springer-Verlag, London.

Zellner, Z. (1971). *An Introduction to Bayesian Inference in Econometrics*. John Wiley and Sons, New York.