Available online at http://scik.org
Commun. Math. Biol. Neurosci. 2026, 2026:52
https://doi.org/10.28919/cmbn/9756
ISSN: 2052-2541

OUTLIER DETECTION AND CONTROL IN BAYESIAN VECTOR AUTOREGRESSIVE PROCESSES

THEOPHILUS ASAMOAH[1,*], ANTHONY GICHUHI WAITITU[2], BISMARK KWAO NKANSAH[3], CYPRIAN OMARI[4]

[1]Department of Mathematics, Institute for Basic Sciences, Technology and Innovation, Pan African University, Juja, Kenya

[2]Department of Statistics and Actuarial Sciences, Jomo Kenyatta University of Agriculture and Technology, Juja, Kenya

[3]Department of Statistics, University of Cape Coast, Cape Coast, Ghana

[4]Department of Statistics and Actuarial Science, Dedan Kimathi University of Technology, Nyeri, Kenya

Abstract: While outlier detection and control are not fundamental to the estimation of Bayesian Vector Autoregressive (BVAR) models, they represent a significant improvement that may enhance the robustness and reliability of estimation and forecasting. The increasing prevalence of big data pose challenges for monitoring and maintaining data quality for estimation. In this paper, the standard BVAR model is extended through Outlier Detection and Control (ODC), known as Extended BVAR-ODC model. The Extended BVAR-ODC decomposes the data generating process into three explicit components: a core VAR process with coefficient matrices and Gaussian innovations, an outlier component, and a multivariate normal-tempered innovation structure. Also, the ODC mechanism operates through a two-stage Bayesian procedure: posterior inference on outlier indicators, and simultaneous estimation of outlier impact coefficients and VAR parameters. It is shown in simulation that the Extended BVAR-ODC provides superior fit compared to the standard BVAR. Across all diagnostics, including parameter recovery, posterior trace convergence, mixing and predictive calibration, the Extended BVAR-ODC consistently outperforms the

*Corresponding author
E-mail addresses: theasamoah36@gmail.com, theophilus.asamoah@students.jkuat.ac.ke
Received December 26, 2025

standard BVAR. Particularly, the posterior estimates of intercepts, lagged coefficients, and covariance matrices exhibit higher stability and smaller posterior dispersion. Additionally, residual analyses and goodness-of-fit confirm that the Extended BVAR-ODC effectively mitigates distortions induced by outliers. All performance measures favour the Extended BVAR-ODC, highlighting its superior generalization performance. Therefore, the Extended BVAR-ODC would be very useful in modeling financial systems that are subject to structural breaks or extreme events. Its outlier-induced specification ensures more robust inference under contaminated data distributions.

Keywords: ARIMA; Bayesian; BVAR; multivariate time series; outlier detection and control.

2020 AMS Subject Classification: 62C10, 62F15, 62H30, 62M10, 68T07, 68T10, 91B84.

1. Introduction

Time series models have gained significant importance in forecasting, and are widely used in various research settings [2]. Autoregressive Moving Average (ARIMA) models have achieved significant success and recognized as standard forecasting tools [1, 2]. In ARIMA models, only correlation between a predicted variable and its historical data is used. As a result, different methods for predicting time series have emerged [3, 4]. A variant of these expanded models are known as Vector Autoregressive (VAR) models. These models are effective in macroeconomics [5, 6]. The development of VAR models is a reaction to the increasing doubt about the *"incredible identifying restrictions"* that affected large-scale macroeconomics models [7]. However, a significant limitation of VAR is the *"curse of dimensionality"* due to high degree of generality and ability to capture complicated dynamic interactions. An alternative to overcoming the over-parameterization problem of VAR is Bayesian estimation. That is, Bayesian VAR (BVAR) models resolve the issue of parameter instability [3]. By including priors, these constraints effectively address the issue. The approach may be seen as an hybrid estimation process. The motivation for the use of BVAR is to obtain accurate estimates and forecasts [3].

In spite of the effectiveness of BVAR , technological advances have facilitated the accumulation of large amounts of data with substantial complexities over time in several domains which have posed considerable challenge to these models. Ensuring data quality is essential for any reliable research. The increasing prevalence of big data poses unique issues for monitoring and maintaining data quality. Literature show tasks such as outlier detection, classification, and clustering [8]. An essential component of data quality monitoring is detecting and controlling

outliers. Statistical models such as conventional VAR and BVAR are well-known for their sensitivity to outliers and lack ability to detect and control outlier effects. Hence, there is the need to incorporate outlier control features into BVAR.

An outlier is a data point that does not follow the normal pattern, or aligns closely with a known aberrant pattern [9]. They may occur due to: malicious behaviours, instrumental mistakes, human or setup errors, environmental changes, and catastrophic events. Irrespective of the source, outliers are interesting to the researcher since they provide distinct information compared to others [10]. Outliers manifest in many ways across different applications. Outliers are put into 3 categories: Category I refers to individual data points that are isolated within a dataset. These are the most basic and easily detectable. In terms of attributes, they are significantly distant from others [11]. Category II usually includes additional contextual factors like time and place [11]. For Category III, a certain subset of the data points exhibit outlier behaviour in relation to the overall data. Individually, none of the data points in a tiny subset deviates significantly from others. However, when considered as a whole, the subset exhibits outlier behaviour. For contextual data like time series, the data is organized as a sequence. Hence, a certain subsequence may be considered an outlier in relation to the full sequence, known as innovation outliers. These are challenging to detect due to their ability to impact others within the same context and attempt to conceal their presence. The control of outliers in time series focuses on aberrant behaviours of the series throughout time [12]. Due to their effect on the estimation of parameters, it is of interest to robustify the standard BVAR to outlier effects.

From literature, numerous challenges associated with outlier detection and control are identified. These challenges range from loss of information and linear relationships to overfitting and computational complexity. In dealing with these challenges, some techniques were proposed. While some focus on outliers in VAR [13], others focused on outliers in BVAR [14, 15]. Some of these studies focused on modeling effects of COVID-19 pandemic related outliers by incorporating a stochastic volatility [14, 15]. It is found that these outliers have significant effects on coefficients without stochastic volatility [16], unless prior for coefficient is adjusted to be highly restrictive. The issue of outliers are controlled using techniques like BVAR with and without stochastic volatilities [14, 15, 16]. This is because earlier work [16] advocate for

stochastic volatility specifications for controlling outlier effects on models other than triangular decomposition. The resultant models are referred to as BVAR with and without stochastic volatility. Also, [17] emphasizes the importance of identifying outliers and structural breaks in time series, as they significantly affect parameter estimation and forecasting accuracy. Further, [18, 17] establishes that statistical properties of time series change at unknown points, and that outliers are conceptually related to but distinct from change points, and that neural networks are valid tools for both detection and estimation in this setting. In this study, the outlier problem is solved through deep learning and density-based outlier detection technique application.

Different from others, [14, 15] obtained parameter estimates by optimizing marginal likelihood to control outliers. In an earlier study, [19] in solving outlier problem, derived an explicit expression for one of the sum of squares cross product matrices involved in outliers, and showed that the discordancy of multiple outliers is preserved along multiple-outlier displaying components with much lower dimensions than the original high-dimensional data. The current study, however, focuses on outlier detection and control concurrently. The explicit outlier detection using the hybrid DNN-DBSCAN technique [20] contrasts with the stochastic volatility method of [21], providing interpretable discrete outlier identification instead of smooth time-varying volatility. Both strategies tackle non-Gaussian innovations through complementary mechanisms. The existing BVAR [22, 23] is enhanced by integrating Minnesota shrinkage with Student-t robustness and structured outlier effects through a diagonal matrix C [24]. To address the challenges of outlier contamination and parameter estimations, a mixture modeling framework that distinguishes between normal and contaminated data is proposed. The BVAR assumes the innovations follow a multivariate normal distribution is define as

$$f(y \mid \mu, \Sigma) = \frac{1}{(2\pi)^{n/2}|\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(y-\mu)^{\top}\Sigma^{-1}(y-\mu)\right) \tag{1}$$

where $y \in \mathbb{R}^n$ is a random vector, $\mu \in \mathbb{R}^n$ is the mean vector, Σ is an $n \times n$ positive definite covariance matrix, n is the dimension of the distribution and π appears in the normalizing constant $(2\pi)^{n/2}$. While this specification provides computational tractability and efficient estimation under Gaussian assumptions, it is highly sensitive to outliers due to exponential decay of the tails [25, 26]. The quadratic component, $(y-\mu)^{\top}\Sigma^{-1}(y-\mu)$ in the exponent assigns negligible

probability to observations far from the mean, causing extreme values to exert disproportionate influence on parameter estimates. This phenomenon is known as the leverage effects [27].

To accommodate the heavy-tailed behaviour of observations and reduce outlier sensitivity, the multivariate Student-t distribution with ν degrees of freedom is incorporated, defined as

$$f(\mathbf{y} \mid \mu, \Sigma, \nu) = \frac{\Gamma\left(\frac{\nu+n}{2}\right)}{\Gamma\left(\frac{\nu}{2}\right)} \frac{1}{(\nu\pi)^{n/2}|\Sigma|^{1/2}} \left[1 + \frac{1}{\nu}(\mathbf{y}-\mu)^{\top}\Sigma^{-1}(\mathbf{y}-\mu)\right]^{-\frac{\nu+n}{2}} \tag{2}$$

where $\mu \in \mathbb{R}^n$ is the location vector, Σ is an $n \times n$ positive definite scale matrix, $\nu > 0$ is the degrees of freedom parameter, and $\Gamma(\cdot)$ is the Gamma function. Unlike the normal distribution, the t-distribution exhibits polynomial rather than exponential tail decay, making it more robust to outliers [28]. The ν controls the tail heaviness. Hence, as ν approaches infinity, the Student-t converges to the normal distribution, as shown in Equation (1), while smaller ν values produce fatter tails that assign substantial probability to extreme observations [29]. This makes the Student-t distribution naturally robust to outliers, as extreme values receive non-negligible likelihood and exert bounded influence on parameter estimates [30]. The critical difference between Equations (1) and (2) lies in the rate of tail decay. The multivariate normal assigns probability proportional to $\exp(-d^2/2)$, resulting in exponentially vanishing probability for large d. In contrast, the Student-t assigns probability proportional to $(\nu)^{-(\nu+n)/2}$, yielding polynomial decay of order $d^{-(\nu+n)}$ [31]. For low $(\nu \approx 5)$, this polynomial decay produces substantially heavier tails. This characteristic enables robust estimation, where outliers contribute to the likelihood without dominating it, and reducing their leverage on coefficient estimates [6].

Following the robust Bayesian framework proposed and extended for time-varying parameter models [28, 32], there is the need to formulate a mixture distribution that combines both specifications. This formulation essentially explicitly separates the data-generating processes, the normal component in Equation (1) governs normal observations arising from the underlying VAR dynamics, while the Student-t component in Equation (2) captures outlying observations [10]. Furthermore, a mixture weight π is estimated from the data, allowing the model to learn the contamination level rather than assuming it a priori, a key advantage over fixed robustification schemes [10]. The challenges of existing methods are addressed by proposing a technique

to control outliers within BVAR through explicit modeling of heavy-tailed innovations. Outliers are first detected by a hybrid DNN-DBSCAN technique [20], and information on outliers, including their locations, magnitudes, and estimated contamination proportion, is incorporated into the standard BVAR to control outlier effects. Specifically, the detected outliers inform the initialization of π and guide the specification of informative priors on ν, ensuring that the Student-t component adequately captures the observed tail behaviour [33, 34]. This paper therefore extends the standard BVAR to obtain what is referred to as the Extended BVAR-ODC model, which jointly estimates AR parameters, covariance structure, mixture weights, and tail heaviness parameters in a coherent Bayesian framework. The resulting model achieves robustness without sacrificing efficiency for clean data, as posterior inference naturally down-weights contaminated observations while preserving full statistical information from regular observations [6, 28]. The rest of the paper is organised as follows: In Section 2, the procedures for Extending the standard BVAR are presented while in Section 3, the main results are presented. Finally, Section 4 presents the conclusion of the paper.

2. Methods

2.1 Extended BVAR-ODC Model Development

In extending the standard BVAR, the hybrid DNN-DBSCAN technique [20] is used to detect outliers. A binary indicator vector is defined for detected outliers, and information on outliers are incorporated into the standard BVAR, as presented in the ensuing sections.

2.1.1 Distribution of clean data

The standard BVAR model for a k-dimensional time series, $\mathbf{Y}_t$ of length T is defined as

$$\mathbf{Y}_t = \underbrace{\mathbf{A}_0 + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{Y}_{t-i}}_{\text{Core process component}} + \underbrace{\varepsilon_t}_{\text{Innovation component}} \tag{3}$$

where, $\mathbf{Y}_t \in \mathbb{R}^{k\times 1}$ is vector of endogenous variables at time t, $\mathbf{A}_0 \in \mathbb{R}^{k\times 1}$ is a vector of intercepts, $\mathbf{A}_i \in \mathbb{R}^{k\times k}$ is coefficient matrix for lag i, $\mathbf{Y}_{t-i} \in \mathbb{R}^{k\times 1}$ is lagged vector of endogenous variables, and $\varepsilon_t \sim N_k(\mathbf{0}_k, \Sigma)$ is Gaussian noise with covariance matrix Σ. Assuming there are no outliers,

the innovation, ε_t in Equation (3) follows the Multivariate Normal Distribution (N_k) defined as

$$\varepsilon_t \sim N_k(\mathbf{0}_k, \Sigma), \quad \text{iid over } t \tag{4}$$

and the conditional distribution becomes

$$\mathbf{Y_t} \mid \mathbf{Y}_{t-1}, \mathbf{Y}_{t-2} \ldots, \mathbf{Y}_{t-p} \sim N_k \left(\mathbf{A}_0 + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{Y}_{t-i}, \Sigma \right) \tag{5}$$

2.1.2 Distribution of contaminated data

If a given multivariate time series data contain outliers, then, a binary indicator variable, $\mathbf{O}_{jt}$ is

$$\mathbf{O}_{jt} = \begin{cases} 1 & \text{If } Y_{jt} \text{ is an outlier} \\ 0 & \text{Otherwise} \end{cases} \tag{6}$$

where O_{jt} is t^{th} outlier at time t for the j^{th} component of the time series vector, Y_{jt} is t^{th} observation at time t for the j^{th} component of the time series vector, $j = 1,2,3,\cdots,k$ and $t = 1,2,3,\cdots,T$. At each time point t, the indicators are collected into a k-dimensional vector as

$$\mathbf{O}_t = (O_{1t}, O_{2t}, \ldots, O_{kt})' \in \{0,1\}^k \tag{7}$$

where the $\mathbf{O}_t$ are treated as fixed and external. Due to outliers, the Gaussian assumptions in Equations (4) is violated. A flexible distribution that accounts for outliers without over-penalizing them is student t-distribution, with a fat tail that downweights outliers naturally [30, 3]. Hence, the outlier component follow the multivariate t-distribution, conditionally defined as

$$\mathbf{Y_t} \mid \mathbf{O}_t \sim t_\nu \left(\nu, \mathbf{A}_0 + \sum_{i=1}^{p} \mathbf{A}_i \mathbf{Y}_{t-i}, \Sigma \right) \tag{8}$$

where, t_ν is t-distribution with ν degrees of freedom, allowing for heavier tails than a normal distribution and captures outliers robustly. The density function is in Equation (2). However, to accommodate both structural outliers and stochastic heavy-tailed behaviour, the innovation distribution is specified as a Mixture of Normal-Student-t (MNT) distribution is

$$\mathbf{E}_t \sim MNT_k(\mathbf{0}_k, \Sigma, \pi, \nu), \quad \textit{i.i.d. over } t \tag{9}$$

with mixture density

$$f(\mathbf{E}_t) = \pi \times \phi_k(\mathbf{E}_t; \mathbf{0}_k, \Sigma) + (1-\pi) \times t_{\nu,k}(\mathbf{E}_t; \mathbf{0}_k, \Sigma) \tag{10}$$

where $\mathbf{E}_t$ is mixture of Gaussian Student-t noise, $\phi_k(\cdot; \mathbf{0}_k, \Sigma)$ is k-variate Normal density with mean zero and covariance Σ, and $t_{\nu,k}(\cdot; \mathbf{0}_k, \Sigma)$ is k-variate Student-t density with ν degrees of freedom, scale matrix is Σ and $\mathbf{C}$ is diagonal. The parameter $\pi \in (0,1)$ is probability of the normal component. Incorporation information on outliers detected in Equation (7) into Equation (3) becomes

$$\mathbf{Y_t} = \underbrace{\mathbf{A_0} + \sum_{i=1}^{p} A_i \mathbf{Y}_{t-i}}_{\text{Core process component}} + \underbrace{\mathbf{CO}_t}_{\text{Outlier component}} + \underbrace{\mathbf{E_t}}_{\text{Innovations component}} \tag{11}$$

where $\mathbf{E}_t$ follows the *MNT* with $\mathbf{C}$ and Σ, with properties of $\mathbf{C}$ as follows

(a) $\mathbf{O}_t$ is binary and by definition

$$E(\mathbf{CO}_t) = \mathbf{C} \times p(\mathbf{O}_t = 1) \tag{12}$$

(b) Typically,

$$|\mathbf{C}| < 1 \quad \text{(additive, mild shocks)}, \tag{13}$$

and may be more than 1 with outlier size.

(c) If k is the number of variables and d is number of outliers detected, then, $\mathbf{O}_t \in \mathbb{R}^{d \times k}$ and $\mathbf{C} \in \mathbb{R}^{d \times k}$

2.1.3 Modified BVAR likelihood function

Now, each Y_{jt} given its history and outliers detected in Equation (6) follow the *MNT* distribution, and the likelihood function of Equation (3) is affected. If the following equations are further defined

$$\mathbf{X}_t = [\mathbf{Y}_{t-1}^\top, \mathbf{Y}_{t-2}^\top, \mathbf{Y}_{t-3}^\top, \ldots, \mathbf{Y}_{t-p}^\top, \mathbf{O_t}^\top]^\top \tag{14}$$

and

$$\Phi = [\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots, \mathbf{A}_p, \mathbf{C}] \in \mathbb{R}^{k \times (kp+k)}, \tag{15}$$

Then, Equation (11) becomes

$$\mathbf{Y} = \mathbf{A}_0 + \Phi \mathbf{X}_t + \mathbf{E}_t \tag{16}$$

where

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1^\top \\ \mathbf{Y}_2^\top \\ \mathbf{Y}_3^\top \\ \vdots \\ \mathbf{Y}_T^\top \end{bmatrix} \quad \text{and} \quad \mathbf{E} = \begin{bmatrix} \mathbf{E}_1^\top \\ \mathbf{E}_2^\top \\ \mathbf{E}_3^\top \\ \vdots \\ \mathbf{E}_T^\top \end{bmatrix}. \tag{17}$$

and the updated likelihood of Equation (16) becomes

$$\begin{aligned} L(\mathbf{A}_0, \Phi, \Sigma \mid \mathbf{Y}, \mathbf{X}, \mathbf{O}) = \prod_{t=1}^{T} \big[\pi \times \Phi_k(\mathbf{Y}_t; \mathbf{A}_0 + \phi X_t, \Sigma) \quad + \\ (1-\pi) \times t_{\nu,k}(\mathbf{Y}_t; \mathbf{A}_0 + \Phi \mathbf{X}_t, \Sigma) \big] \end{aligned} \tag{18}$$

where, $\phi_k(\cdot; \mu, \Sigma)$ is density of N_k, $t_{\nu,k}(\cdot; \mu, \Sigma)$ is density of a k-variate t-distribution with ν degrees of freedom.

2.1.4 Prior distributions of parameters

Assuming Bayesian setup, the priors are specified to capture the outlier effect coefficients, $\mathbf{C}$ as follows

$$\begin{cases} \mathbf{A_0} \sim N_k(m_0, V_0) \\ \Phi \sim MN_{k\times(kp+k)}(M_\Phi, V_\Phi, \Sigma) \\ \Sigma \sim IW(\Psi, \nu_0) \\ \nu \sim \text{Gamma}(a_\nu, b_\nu) \\ \pi \sim \text{Beta}(a_\pi, b_\pi) \end{cases} \tag{19}$$

where $MN_{k\times d}(M, V, \Sigma)$ is the matrix-normal distribution with dimensions $k \times d$, M_Φ is the prior mean matrix for Φ, V_Φ is column covariance matrix, and Σ is row covariance matrix. Also, $IW(\Psi, \nu_0)$ is the IW distribution, $\Psi \in \mathbb{R}^{k\times k}$ is a positive-definite scale matrix, and $\nu_0 > k-1$ is df. Specifically, a diffuse Gaussian prior with substantial variance is used for the intercept

vector, causing it to be weakly informative and allowing the data to take control of the posterior [35], while also guaranteeing the proper presence of the posterior [36]. The zero-centered prior is suitable following the data normalization, as intercepts denote departures from standardized means. This prevents the imposition of prior beliefs about the nature of economic variables, which may exhibit non-stationarity [37]. The diffuse prior guarantees robustness against prior misestimation [38]. The matrix-normal prior with Minnesota-type shrinkage is the accepted standard for BVAR models, offering automatic regularization for the high-dimensional parameter space [39]. The Gaussian assumption for Φ yields exponential tail decay, which is suitable considering that VAR coefficients are constrained in stable systems. This minimizes inflated coefficient estimates that could contradict stationarity [40]. The IW prior is conjugate to the Gaussian likelihood, facilitating efficient Gibbs sampling and guaranteeing positive-definiteness of Σ [41], making it highly relevant for current BVAR applications due to its computational efficiency [42]. Primarily, ν_0 is set up in order to ensure a finite prior mean, consequently giving a weakly informative prior centered around the identity. This indicates unit unconditional deviations following standardization [43]. The identity scale matrix $\Psi = I_k$ imposes minimal prior information on the covariance structure, permitting the data to ascertain correlations across variables [38]. Furthermore, the gamma prior on ν permits data-driven adjustment of tail heaviness, combining robustness against outliers with efficiency under normal conditions [24]. This is essential for economic data demonstrating extreme events [44]. The gamma distribution features an exponential right tail, allowing for arbitrarily significant ν while assigning minimal prior probability to extremely heavy tails. This indicates a prior assumption that extreme Cauchy-like behaviour is rare but feasible [45]. The beta prior on π regulates the balance between the Gaussian and the Student-t components in the innovation distribution. The symmetry provides minimal regularization to uniform weighting while remaining weakly informative [46]. This reduced label switching and stabilized the MCMC sampling, enabling the data to ascertain the dominant component [46]. Consequently, these priors are weakly informative, enabling the data to prevail in the posterior while offering regularization for stable estimate.

2.1.5 Posterior distributions

Due to the structure of Equation (16), the posterior distribution of the priors in Equation (19) are not analytically tractable, but holds partially for Σ, Φ and $\mathbf{A_0}$. Hence, the posteriors are specified up to proportionality for each parameter as follows

$$
\begin{cases}
p(\mathbf{A}_0 \mid \Phi,\Sigma,\nu,\pi,\mathbf{Y},\mathbf{X},\mathbf{O}) & \propto \prod_{t=1}^{T} \Big[\pi \times \phi_k(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)+(1-\pi) \quad \times \\
& \qquad t_{\nu,k}(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)\Big] \times N_k(\mathbf{A}_0 \mid m_0,V_0) \\
p(\Phi \mid \mathbf{A}_0,\Sigma,\nu,\pi,\mathbf{Y},\mathbf{O}) & \propto \prod_{t=1}^{T} \Big[\pi \times \phi_k(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)+(1-\pi) \quad \times \\
& \qquad t_{\nu,k}(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)\Big] \times MN_{k\times(kp+k)}(\Phi \mid M_\Phi,V_\Phi,\Sigma) \\
p(\Sigma \mid \mathbf{A}_0,\Phi,\nu,\pi,\mathbf{Y},\mathbf{X},\mathbf{O}) & \propto \prod_{t=1}^{T} \Big[\pi \times \phi_k(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)+(1-\pi) \quad \times \\
& \qquad t_{\nu,k}(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)\Big] \times IW(\Sigma \mid \Psi,\nu_0) \\
p(\nu \mid \mathbf{A}_0,\Phi,\Sigma,\pi,\mathbf{Y},\mathbf{X},\mathbf{O}) & \propto \prod_{t=1}^{T} \Big[\pi \times \phi_k(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)+(1-\pi) \quad \times \\
& \qquad t_{\nu,k}(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)\Big] \times \text{Gamma}(\nu \mid a_\nu,b_\nu) \\
p(\pi \mid \mathbf{A}_0,\Phi,\Sigma,\nu,\mathbf{Y},\mathbf{X},\mathbf{O}) & \propto \prod_{t=1}^{T} \Big[\pi \times \phi_k(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)+(1-\pi) \quad \times \\
& \qquad t_{\nu,k}(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)\Big] \times \text{Beta}(\pi \mid a_\pi,b_\pi)
\end{cases} \tag{20}
$$

Therefore, the joint posterior distribution is defined as

$$
\begin{aligned}
p(\mathbf{A}_0,\Phi,\Sigma,\nu,\pi \mid \mathbf{Y},\mathbf{X},\mathbf{O}) \propto & \underbrace{L(\mathbf{Y} \mid \mathbf{A}_0,\Phi,\Sigma,\nu,\pi,\mathbf{X},\mathbf{O})}_{\text{Mixture Likelihood Component}} \times \\
& \underbrace{p(\mathbf{A}_0) \times p(\Phi) \times p(\Sigma) \times p(\nu) \times p(\pi)}_{\text{Prior Distributions Component}}
\end{aligned} \tag{21}
$$

and substituting Equations (19) and (20) into Equation (21), and Equation (21) becomes

$$
\begin{aligned}
p(\Phi,\mathbf{A}_0,\Sigma,\nu,\pi \mid \mathbf{Y},\mathbf{X},\mathbf{O}) \propto & \prod_{t=1}^{T} \Big[\pi\,\phi_k(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)+(1-\pi)\,t_{\nu,k}(\mathbf{Y}_t;\mathbf{A}_0+\Phi\mathbf{X}_t,\Sigma)\Big] \\
& \times MNT_k(\Phi \mid M_\Phi,V_\Phi,\Sigma)\,N_k(\mathbf{A}_0 \mid m_0,V_0)
\end{aligned} \tag{22}
$$

$$\times IW(\Sigma \mid \Psi, \nu_0)\,\text{Gamma}(\nu \mid a_\nu, b_\nu)\,\text{Beta}(\pi \mid a_\pi, b_\pi)$$

2.1.6 Extended BVAR-ODC

The adjusted predictions from the Extended BVAR-ODC accounts for $\mathbf{C}$, and defined as

$$\widehat{\mathbf{Y}}_t = \widehat{\mathbf{A}}_0 + \sum_{i=1}^{p} \widehat{\mathbf{A}}_i \mathbf{Y}_{t-i} + \widehat{\mathbf{C}}\mathbf{O}_t \tag{23}$$

Equation (23) is to correct distortions introduced by outliers, leading to a realistic residual structure that yields enhanced robustness for structural inference under contaminated data.

2.2 Modeling Design and Assumptions

The VAR coefficient matrices are computed using Bayesian inference rather than being predetermined [47]. Also, Σ is considered as an unknown quantity that is evaluated concurrently with the VAR coefficients. To ensure parsimony and computational efficiency, an assumption is made regarding the time-invariance of Σ, while recognizing that this may be compromised in the presence of outliers. The Inverse-Wishart prior on Σ is conjugate to the Gaussian likelihood, enabling closed-form posterior updates. Outliers are directly represented by $\mathbf{CO}_t$ instead of being included into time-varying volatility, facilitating clearer understanding. It is also assumed that outliers are unusual occurrences, as evidenced by the examination of standard economic time series in which rare events are observed. The framework incorporates additive and innovative outliers by means of appropriate specification of $\mathbf{C}$. Moreover, outliers provide significant divergence from the predicted values, allowing posterior probabilities $P(\mathbf{O}_t = 1|data)$ to differentiate between outlying and non-outlying observations. Σ is a $k \times k$ symmetric, positive-definite matrix that represents contemporaneous correlations among the k series. Finally, Σ remains constant over time. To provide robustness while retaining conjugacy features, we utilize the MNT distribution for $\mathbf{E}_t$, a minor generalization of the Gaussian that allows for regulated heavy-tailedness.

2.3 MCMC Parameter Estimations Procedures

Posterior inference for the Extended BVAR-ODC is obtained by a Gibbs sampling, incorporating auxiliary phases for non-conjugate parameters. Each iteration of the MCMC occurs as follows: The coefficient matrix for VAR coefficients is sampled based on the data and current outlier indicators, applying the Minnesota prior, with the conditional posterior being Gaussian.

Furthermore, for Σ, it is drawn from its inverse-Wishart posterior distribution. Also, for each observation at t, the binary indicator $\mathbf{O}_t$ is sampled from a Bernoulli distribution with probability π. Moreover, given that $\mathbf{O}_t = 1$, $\mathbf{C}$ is drawn from its Gaussian posterior distribution. Also, the outlier probability parameter π is revised based on its Beta posterior distribution. The tail parameter ν, which regulates the outlier component, is finally updated by a Metropolis–Hastings step due to a conjugate posterior absence. The algorithm runs for N iterations, ignoring an initial burn-in phase. Posterior summaries are derived from the retained draws.

2.4 Performance Measures

The performance of the Extended BVAR-ODC is evaluated in relation to the standard BVAR. The stability and accuracy of the estimations are assessed to compare the estimates between the two models. For a 2-variable time series, the Extended BVAR-ODC at lag 1 is

$$\mathbf{Y}_t = \mathbf{A}_0 + \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{C}\mathbf{O}_t + \mathbf{E}_t, \quad \mathbf{E}_t \sim MNT_k(\mathbf{0}_k, \Sigma, \pi, \nu) \tag{24}$$

and that for the Standard BVAR is

$$\mathbf{Y}_t = \mathbf{A}_0 + \mathbf{A}_1 \mathbf{Y}_{t-1} + \varepsilon_t, \quad \varepsilon_t \sim N_k(\mathbf{0}_k, \Sigma) \tag{25}$$

where k is number of variables. The true parameters, estimated parameters are presented in Table 1, together with other important performance metrics.

2.4.1 Posterior Predictive Check

Posterior Predictive Check assesses how well a model reproduces key features of the observed data [48, 49]. Given the observed data $\mathbf{D} = \{\mathbf{Y}_t\}_{t=1}^{T}$ and the posterior samples of parameters, θ, the posterior predictive distribution is

$$p(\mathbf{Y}^{\text{rep}} \mid \mathbf{D}) = \int p(\mathbf{Y}^{\text{rep}} \mid \theta)\, p(\theta \mid \mathbf{D})\, d\theta \tag{26}$$

In Equation (11), the posterior predictive samples $\mathbf{Y}_t^{\text{rep}}$ are drawn by simulating from Equation (26) using the posterior draws of the parameters, $(\mathbf{A}_i, \mathbf{C}, \Sigma_{ij})$ of the Extended BVAR-ODC.

2.4.2 Probabilistic Evaluation Metrics

Two probabilistic evaluation metrics, namely, Probability Integral Transform (PIT) and Continuous Ranked Probability Score (CRPS) are applied, as presented in the ensuing sections:

(a) PIT verifies consistency of predictive distribution with observed data [50], as

$$\text{PIT}_t = F_t(\mathbf{Y_t}) \tag{27}$$

where $F_t(\cdot)$ denotes cumulative distribution function of posterior predictive distribution of the new model at t, on $U(0,1)$. For the PIT technique, deviations from uniformity indicate systematic biases in predictions of the Extended BVAR-ODC [49].

(b) CRPS provides a scalar summary of predictive accuracy by comparing entire predictive distribution to observed outcome. CRPS for a predictive CDF F and $\mathbf{Y}$ is

$$\text{CRPS}(F,Y) = \int_{-\infty}^{\infty} [F(Z) - \mathbb{I}\{Y \leq Z\}]^2 \, dZ, \tag{28}$$

where $\mathbb{I}\{\cdot\}$ is an indicator function. Smaller CRPS values indicate better predictive sharpness and calibration [48].

2.4.3 Goodness-of-Fit Metrics

The predictive performances are assessed using Mean Absolute Error (MAE), Root Mean Square Error $(RMSE)$, Mean Square Error (MSE), Mean Absolute Deviations (MAD) and Coefficient of Determination (R^2) as presented as follow

$$\begin{cases} \text{MSE} = \frac{1}{TK}\sum_{t=1}^{T}\sum_{k=1}^{K}(\mathbf{Y}_{tk} - \hat{\mathbf{Y}}_{tk})^2 \\ \text{RMSE} = \sqrt{\frac{1}{TK}\sum_{t=1}^{T}\sum_{k=1}^{K}(\mathbf{Y}_{tk} - \hat{\mathbf{Y}}_{tk})^2} \\ \text{MAE} = \frac{1}{TK}\sum_{t=1}^{T}\sum_{k=1}^{K}|Y_{tk} - \hat{\mathbf{Y}}_{tk}| \\ \text{MAD} = \frac{1}{TK}\sum_{t=1}^{T}\sum_{k=1}^{K}|\mathbf{Y}_{tk} - \bar{\mathbf{Y}}_{k}| \\ R^2 = 1 - \dfrac{\sum_{t=1}^{T}\sum_{k=1}^{K}(\mathbf{Y}_{tk} - \hat{\mathbf{Y}}_{tk})^2}{\sum_{t=1}^{T}\sum_{k=1}^{K}(\mathbf{Y}_{tk} - \bar{\mathbf{Y}}_{k})^2} \end{cases} \tag{29}$$

where where T represents the number of observations and K is the number of variables. Hence, lower values of the metrics imply better model fit and vice versa.

2.4.4 Data Generation Process

A bivariate time series of $T = 1{,}000$ daily observations, representing exchange rate and inflation over approximately 3 years (January 2023 – December 2025), is simulated to evaluate the Extended BVAR-ODC. The sample size is sufficient for stable posterior estimation and meaningful evaluation under controlled simulation conditions. Parameters follow recommendations of [53], [52], and [51], with the DGP constructed to mirror the dual-robustness mechanism of the Extended BVAR-ODC. The multivariate time series $\{\mathbf{Y}_t\}_{t=1}^{T}$ is generated as

$$\mathbf{Y}_t = \mathbf{A}_0 + \mathbf{A}_1 \mathbf{Y}_{t-1} + \mathbf{C}\mathbf{O}_t + \mathbf{E}_t \tag{30}$$

where structural outliers are introduced by randomly selecting $\lfloor 0.10 \times T \rfloor$ time indices from $\{p+1, \ldots, T\}$, setting the corresponding elements of $\mathbf{O}_t \in \{0,1\}^k$ to 1 for randomly activated variables at each selected outlier time and $\mathbf{O}_t = \mathbf{0}_k$ otherwise, inducing additive level shifts representing structural breaks. $\mathbf{C}$ is defined as diagonal so that simulated outlier shocks affect each variable independently, isolating the outlier signal per variable and providing a clean controlled environment for evaluating the DNN-DBSCAN detection step and the Extended BVAR-ODC recovery performance. The 10% outlier proportion used exceeds the conventional 5% acceptable threshold, to rigorously stress-test both models. To mirror the dual-robustness mechanism of the Extended BVAR-ODC, $\mathbf{E}_t$ follows the Mixture Normal-Student-t (MNT) distribution

$$\mathbf{E}_t \sim MNT_k(\mathbf{0}_k, \Sigma, \pi, \nu), \quad \text{i.i.d. over } t \tag{31}$$

with mixture density

$$f(\mathbf{E}_t) = \pi \cdot \phi_k(\mathbf{E}_t; \mathbf{0}_k, \Sigma) + (1-\pi) \cdot t_{\nu,k}(\mathbf{E}_t; \mathbf{0}_k, \Sigma) \tag{32}$$

where $\phi_k(\cdot; \mathbf{0}_k, \Sigma)$ is the k-variate Normal density with mean zero and covariance Σ, $t_{\nu,k}(\cdot; \mathbf{0}_k, \Sigma)$ is the k-variate Student-t density with ν degrees of freedom and scale matrix Σ, and $\pi \in (0,1)$ is the mixing weight on the Normal component, providing continuous heavy-tail robustness to capture innovation outliers distinct from persistent structural breaks. The vector and matrix specifications are: $\mathbf{Y}_t \in \mathbb{R}^{2\times 1}$, $\mathbf{A}_0 \in \mathbb{R}^{2\times 1}$, $\mathbf{A} \in \mathbb{R}^{2\times 2}$, $\mathbf{C} \in \mathbb{R}^{2\times 2}$, $\mathbf{O}_t \in \{0,1\}^{2\times 1}$, and $\Sigma \in \mathbb{R}^{2\times 2}$.

3. Main Results

3.1 Extended BVAR-ODC Parameters Estimation

The Extended BVAR-ODC builds on the standard BVAR by incorporating outliers detected through hybrid DNN-DBSCAN [20], as exogenous binary indicators into the model. Standard BVAR models are effective for multivariate forecasting but produce biased estimates when outliers are present and unaccounted for. By explicitly flagging outliers, the Extended BVAR-ODC separates systematic variation from exceptional deviations, preventing distortion of coefficient estimates. This two-stage approach combines unsupervised machine learning for precise outlier identification with Bayesian econometric methods for stable, accurate parameter estimation.

3.1.1 Intercept terms: Bias and recovery of the intercepts

From Table 1, A_{01} for BVAR exhibits severe overestimation bias, showing 122% positive bias, suggesting that the unmodeled outliers artificially inflate the constant terms [9]. In contrast, the Extended BVAR-ODC recovers with a 42% smaller standard errors, demonstrating superior precision. This shows outliers inflate correlation estimates and distort dependence structure [27]. The inflated estimate of BVAR reflects contamination bias in high-dimensional data streams [10]. Again, for A_{02}, BVAR fails to identify this parameter effectively, showing 93% attenuation to zero. This represents the classic suppression effects where outliers obscure true negative relationships [30]. The Extended BVAR-ODC recovers the true value, confirming that robust modeling through the heavy-tailed distribution restores parameter identifiability [5].

3.1.2 Lagged coefficients: Robustness and efficiency gains

From Table 1, concerning A_{11}, both models perform adequately, but the Extended BVAR-ODC exhibit superior efficiency: 44% smaller standard error and 85% higher t-statistics. The tighter credible interval of the Extended BVAR-ODC reflect effective outlier down-weighting, which is consistent with the robust M-estimation principles [3]. The accuracy improvement confirms that the heavy-tailed error distributions reduce the bias in the dominant autoregressive parameters [6]. For A_{12}, the Standard BVAR shows minimal bias but 73% larger standard error. The Extended BVAR-ODC exhibits slight underestimation (14% negative bias) but

maintains stronger statistical evidence. This trade-off suggests that the robust model conservatively shrinks the weak cross-variable effects to prevent outlier-driven spurious relationships [1]. Also, on A_{21}, the Standard BVAR underestimates by 17%, while the Extended BVAR-ODC achieves 3% accuracy with 36% smaller standard error. The t-statistic of the Extended model nearly doubles, providing substantially stronger evidence for this cross-lag effect [4]. Finally, for A_{22}, although both models show moderate accuracy, the Extended BVAR-ODC reduces bias by about 63% and improves precision by about 33%. The t-statistic also increases by 45%, indicating a more reliable inference for this second-equation autoregressive effect [2].

TABLE 1. Estimated Parameters: Standard vs Extended BVAR-ODC

Parameter	**True value**	**STD BVAR (±SD)**	**t-value**	**p-value**	**Extended BVAR-ODC (±SD)**	**t-value**	**p-value**
A_{01}	0.500	1.111 ± 0.113	9.832	0.000	0.510 ± 0.066	7.727	0.000
A_{02}	-0.500	-0.036 ± 0.099	-0.364	0.716	-0.500 ± 0.066	-7.576	0.000
A_{11}	0.700	0.678 ± 0.032	21.188	0.000	0.705 ± 0.018	39.167	0.000
A_{12}	0.200	0.202 ± 0.045	4.489	0.000	0.172 ± 0.026	6.615	0.000
A_{21}	0.100	0.083 ± 0.028	2.964	0.000	0.103 ± 0.018	5.722	0.000
A_{22}	0.600	0.588 ± 0.039	15.077	0.000	0.568 ± 0.026	21.846	0.000
Σ_{11}	1.000	3.539 ± 0.163	21.712	0.000	1.027 ± 0.023	20.040	0.000
Σ_{12}	0.300	2.243 ± 0.121	18.537	0.000	0.302 ± 0.015	8.343	0.000
Σ_{21}	0.300	2.43 ± 0.121	18.537	0.000	0.302 ± 0.015	8.343	0.000
Σ_{22}	1.000	2.618 ± 0.113	23.168	0.000	1.018 ± 0.022	20.245	0.000
C_{11}	2.000	-	-	-	1.257 ± 0.066	19.046	0.000
C_{22}	1.500	-	-	-	1.380 ± 0.067	20.597	0.000
ν	3.000	-	-	-	3.924 ± 1.176	3.338	0.000
π	0.900	-	-	-	0.880 ± 0.069	12.754	0.000

3.1.3 Covariance parameters: Bias control and true uncertainty

From Table 1, Σ_{11} for BVAR shows overestimation by 254%, the most severe bias. However, the Extended BVAR-ODC achieves high recovery (3% deviation) with 86% smaller standard error, demonstrating that the mixture distributions correctly separate the outlier-induced variance from inherent volatility [30]. For the off-diagonal elements, the Standard BVAR overestimates the covariance by about 647-710%, creating spurious cross-variable dependencies. The Extended BVAR-ODC on the other hand recovers the true covariance with about 1% accuracy and 87% smaller standard error. This demonstrates that robust modeling prevents outliers from inflating correlation structures [5]. Lastly, the Standard BVAR overestimates Σ_{22} by 162%, while the Extended BVAR-ODC achieves 2% accuracy with 81% smaller standard error [3].

3.1.4 Outlier effect coefficients: Novelty in structural parameters

From Table 1, both C_{11} and C_{22} exhibit moderate negative bias: C_{11} is underestimated by 37%, while C_{22} shows substantially better recovery with only 8% underestimation. The moderate underestimation of C_{11} reflects the conservative approach of the model to outlier attribution. This negative bias is characteristic of robust estimation procedures, which deliberately set a high threshold for labeling observations as outliers to avoid false positives [9]. The substantially better recovery of C_{22} is noteworthy. This may arise from several factors. The interaction between hybrid DNN-DBSCAN and Bayesian procedure may favour certain configurations [34].

3.1.5 Mixture distribution parameters: Modelling of tail behaviour

Furthermore, the mixture distribution parameters characterize heavy-tailed error structure. The degrees of freedom is overestimated at 31%. This confirms substantial tail heaviness, implying the Student-t distribution has approximately about $3-4$ times heavier tails than Gaussian distribution [1]. This validates the necessity of flexible error distributions for robust estimations [2]. Finally, π is underestimated at about 2%. Approximately 88% of the data are treated as "clean" while 12% receive outlier-robust treatment. The tight confidence interval suggests that the model confidently distinguishes contaminated data from clean data [9], which is consistent with contamination-aware density estimation [54]. This suggests the model is identifying and

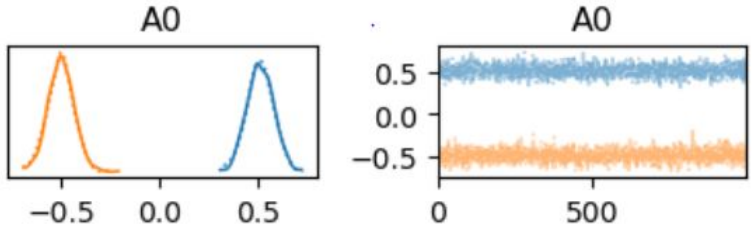
classifying outliers, with benefit of robust parameter estimation, protection against misspecification, and better tail risk modeling. This is because slightly over-conservative outlier detection improves overall model performance in high-dimensional data streams [10].

3.2 Extended BVAR-ODC Performance

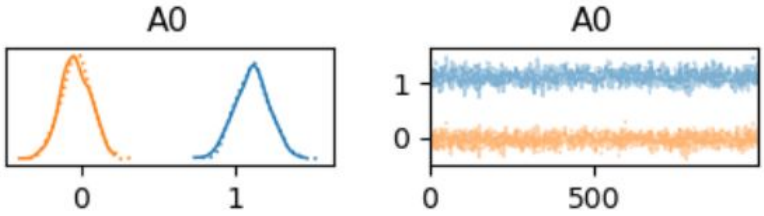
The performance of the Extended BVAR-ODC is evaluated in relation to one of its main variant, the Standard BVAR, as presented in the following sections.

3.2.1 Trace plots and posterior densities

In Figure 1, whereas the Standard BVAR exhibits biased and more dispersed posteriors, the Extended BVAR-ODC recovers the intercept tightly about values near the data generating process. This indicates that outliers inflate intercept estimates when they are not explicitly modeled. The mixture-based likelihoods and outlier-aware preprocessing have reduced the intercept bias and produced better calibrated baseline levels [55]. In Figure 2, both specifications re-



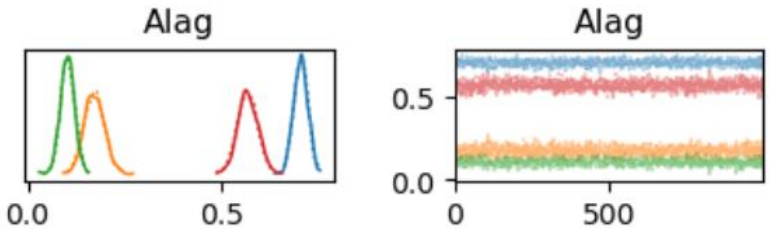
(A) Extended BVAR-ODC Model



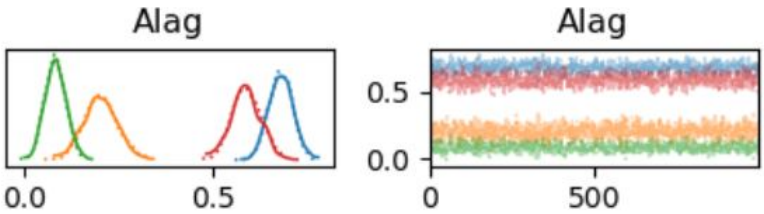
(B) Standard BVAR Model

FIGURE 1. Trace and Posterior Density Plots of Intercepts

cover the signs and dynamic propagation, but the Extended BVAR-ODC appears to yield noticeably narrower posteriors and slightly closer point estimates to the known data generation process for several of the entries. The narrower credible interval indicates improved estimator



(A) Extended BVAR-ODC Model

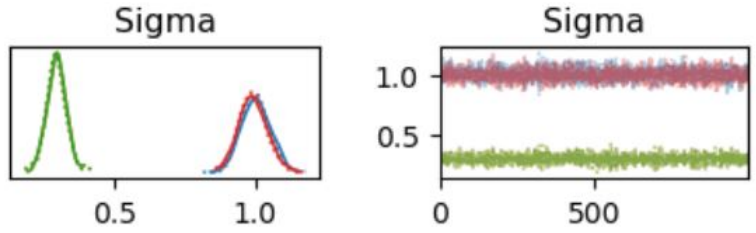


(B) Standard BVAR Model

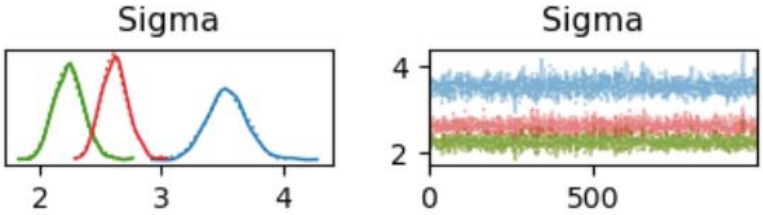
FIGURE 2. Trace and Posterior Density Plots of Lagged Coefficients

efficiency and reduced influence of aberrant observations. Recent advances in BVAR priors and robust likelihoods demonstrate similar gains in coefficient recovery and forecasting performance [53, 56].

Figure 3 highlights pronounced difference. The Standard BVAR produces inflated marginal variances and covariances while the Extended BVAR-ODC recovers values close to the data generation process. This illustrates a well documented phenomenon that, unmodeled outliers can severely bias sample and posterior covariance estimates and create spurious cross-correlations [55, 57]. This means using robust covariance estimation through the Normal-Student-t mixtures distribution reduces the inflation and yields more defensible structural inferences like comovement, shock transmission, among others. The Extended BVAR-ODC indicates meaningful departures from Gaussianity and quantify relative prevalence of atypical observations. The outlier effects matrix, C isolates outlier impacts and explains much variance inflation as observed under the Standard BVAR model. Literature emphasizes that reporting these parameters is very essential [58, 59].



(A) Extended BVAR-ODC Model

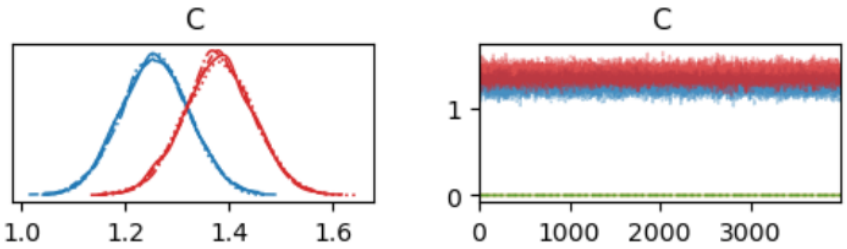


(B) Standard BVAR Model

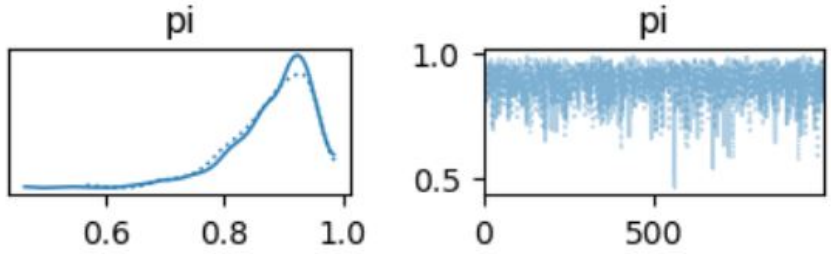
FIGURE 3. Trace and Posterior Density Plots of Sigma Coefficients

On a whole, the trace plots (Figures 1-3) demonstrate good stationarity and mixing for the Extended BVAR-ODC parameters, such as the absence of visible drift, $\hat{R} \approx 1$ and adequate Effective Sample Size, which implies reliable posterior summaries. Although the standard BVAR traces are stable, the posterior uncertainty is inflated for the scale parameters, consistent with available identifiability and efficiency losses when no outlier control (model misspecification) is present [53, 60]. The Extended BVAR-ODC achieves tighter variance estimation and improved outlier control, indicating successful outlier adjustment and enhanced credibility. It is clear that the Extended BVAR-ODC outperforms in all. Its posteriors are tighter, better centered, and more stable, a strong evidence of improved calibration, efficiency, and robustness to outliers.

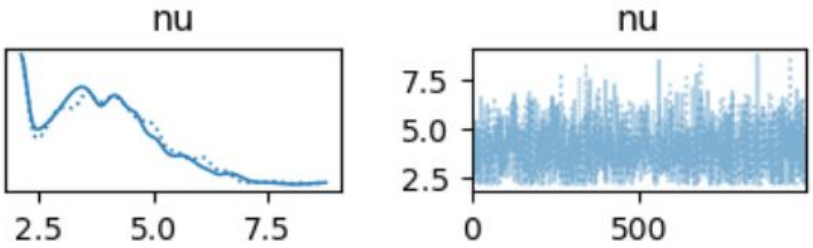
Additionally, in Figure 4, the posterior densities of Figure 4(a) displays a well-formed, concentrated modes and show stable mixing across all chains. The statistically significant, non-zero posterior mass for $\mathbf{C}$ indicate that feature-identified outlier effects materially shift the conditional mean. These coefficients act as explicit correction terms absorbing transient shocks which would otherwise bias the BVAR intercepts and inflate the error covariance. Recent works that combine outlier detection with Bayesian estimation produce interpretable adjustment coefficients and improve parameter recovery and forecast calibration [59, 55]. The π directly informs the degree of heavy-tail behaviour and models effective complexity [58, 57]. For Figure 4(b),



(A) **C**



(B) π



(C) ν

FIGURE 4. Trace and Posterior Density Plots of **C** and Mixture Parameters

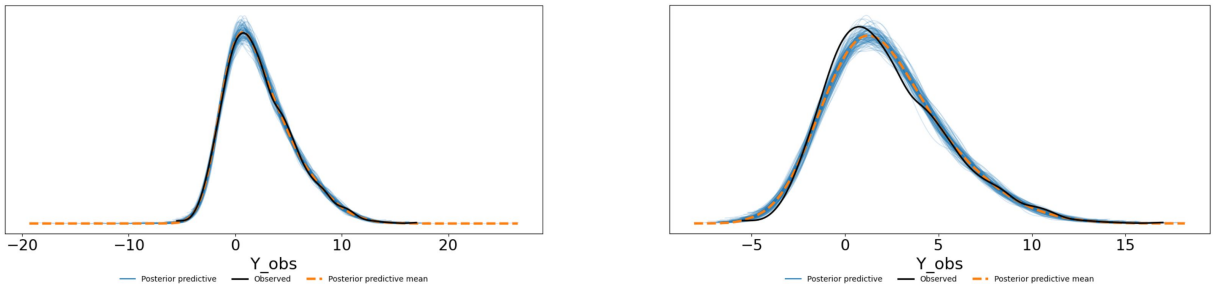
the posterior density and trace concentrate near high values (close to 1), indicating that most of the observations are generated normally while a smaller percentage is handled by the outlier component. This quantifies the prevalence of atypical observations and is important for assessing robustness and model fit. That is, a high posterior π with non-zero C terms implying impactful outliers.

Finally, in Figure 4(c), the posterior for ν shows mass concentrated around low integer values, with evidence of stationarity. The posterior ν signals pronounced heavy-tails and supports the use of mixture likelihoods instead of Gaussian. Learning ν from data improves predictive calibration and reduces variance inflation that otherwise occurs under Gaussian misspecification [57, 55]. The estimated outlier percentage (12%) is consistent with typical contamination rates

in time series data and analysis [10], validating that contamination is common in real-world applications. Also, the tight posterior suggests the Extended BVAR-ODC effectively discriminates between clean data and outliers, improving robustness. The degrees of freedom provides a strong evidence of leptokurtic behaviour in the data, which justifies the use of the mixture distribution over purely Gaussian. For this reason, according to [61], BVAR models with t-distributed errors substantially outperform Gaussian models when $\nu < 10$. Similarly, [34] shows that by adding heavy-tailed distributions in time series models improves outlier detection and control by down-weighting outlying observations appropriately.

3.2.2 Posterior predictive checks

Another important diagnostic tool is PPC, Equation (26). In Figure 5, the orange-dashed lines align closely with the observed central peak in the Extended BVAR-ODC, indicating accurate recovery of the conditional mean. Contrastingly, the Standard BVAR predictive mean tracks the centre but with less concentration of mass at the mode. This shows that the outlier adjustment and heavy-tailed component sharpened mean estimation and reduced bias from extreme observations, consistent with current studies that the Student–t or mixture likelihoods improve point and interval forecasts in heavy-tailed time series [57, 53]. In Figure 5(a), the Extended BVAR-



(A) Extended BVAR-ODC (B) Standard BVAR

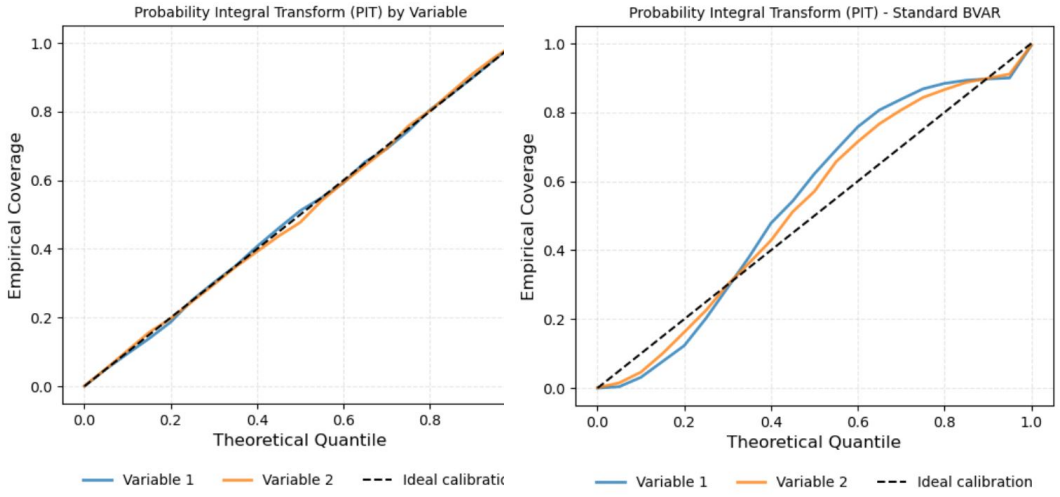
FIGURE 5. Comparison of PPC Plots of Extended BVAR-ODC and BVAR

ODC predictive bands appear noticeably sharper at the mode and better match the observed tail decay. Standard BVAR however shows a wider predictive spread and heavier right-hand tails, which indicates the variance of the Standard BVAR model is inflated by extreme events not modeled explicitly, producing over-dispersed forecasts. Robust likelihoods and explicit outlier

components reduce this inflation and yield more realistic tail coverage, improving probabilistic calibrations (PIT and CRPS) and policy-relevant risk assessment [55, 58].

3.2.3 Probabilistic evaluation

Focusing on model fit, outlier robustness, and predictive reliability, the Extended BVAR-ODC provides a tighter and more accurate posterior predictive fit than the Standard BVAR. On outlier robustness, the inclusion of the outlier suppression term $\mathbf{CO}_t$ and mixture likelihood improve robustness against heavy-tailed outliers. Concerning predictive reliability, the narrower



(A) Extended BVAR-ODC (B) Standard BVAR

FIGURE 6. PIT Plots of Extended BVAR-ODC and Standard BVAR

predictive bands, the Extended BVAR-ODC indicates lower posterior uncertainty and improved calibration. It clearly outperforms the Standard BVAR in terms of fit accuracy, stability, and robustness to outliers, confirming the benefit of outlier control. Hence, the Extended BVAR-ODC effectively captures both central and extreme behaviour, confirming its robustness against outliers. The close overlay of posterior predictive draws with observed curve in the Extended BVAR-ODC panel suggests good probabilistic calibration. The model generates predictive distributions that resemble the empirical distribution across most quantiles. The predictive draws of the Standard BVAR systematically under-represents concentration near the mode and over-represents mass in the tails. This implies miscalibration that may degrade scoring rules and decision performance. In furtherance to this, modern diagnostics tools like PIT and CRPS that

TABLE 2. Forecast Performance: Standard BVAR vs Extended BVAR-ODC

Model Performance Measures	Y_1	Y_2	**Mean / Overall**	**Coverage Target**
CRPS; lower is better				
Standard BVAR	2.0540	1.7852	1.9196	–
Extended BVAR-ODC	**1.1676**	**1.1647**	**1.1662**	–
Coverage Rate (%; target = 95%)				
Standard BVAR	91.89%	92.29%	92.09%	95%
Extended BVAR-ODC	**95.30%**	**95.40%**	**95.35%**	95%

quantify these differences [57, 60], are in Figure 6 and Table 2. It is seen that the Extended BVAR-ODC achieves substantially lower CRPS across variables, indicating sharper and more accurate predictive distributions. On calibration, the Standard BVAR undercovers 92%, meaning its predictive intervals are narrow. However, the Extended BVAR-ODC achieves coverage at the 95% target, indicating well-calibrated uncertainty estimates.

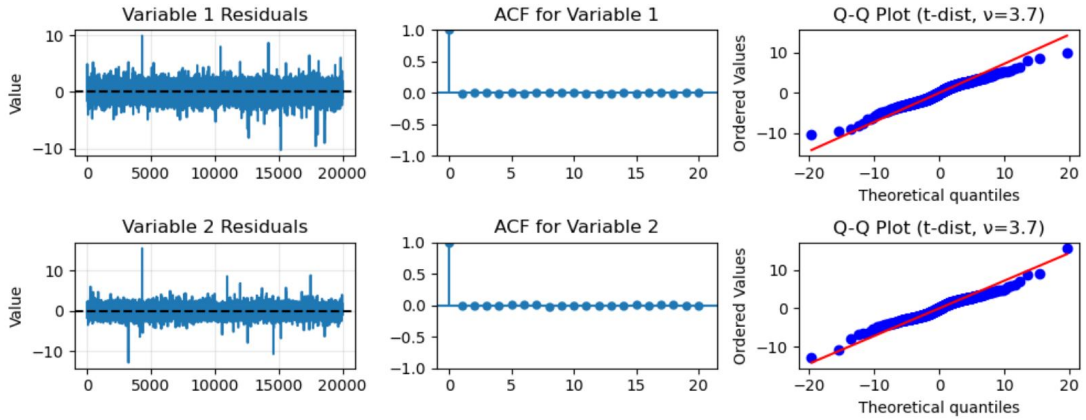
Overall, the Extended BVAR-ODC outperforms the BVAR in accuracy and calibration. This suggests the extensions provided a meaningful gains in posterior predictive reliability. The improvement in the Extended BVAR-ODC reflects the combined effect of the two modeling choices. The mixture model permitted heavy-tail and outlier-effect term preprocessing the hybrid technique to detect outliers. Hybrid approaches combining machine learning oriented outlier detection with Bayesian inference produce similar gains in predictive sharpness and interpretability [59, 55]. Applicably, the Extended BVAR-ODC delivers narrower predictive intervals and improved tail coverage, which translates to better calibrated risk measures. This means a BVAR in the presence of even a few outliers can lead to over-conservative forecasts and misleading assessments of cross-variable comovement [55, 56], supported by Table 2.

3.3 Model Diagnostics

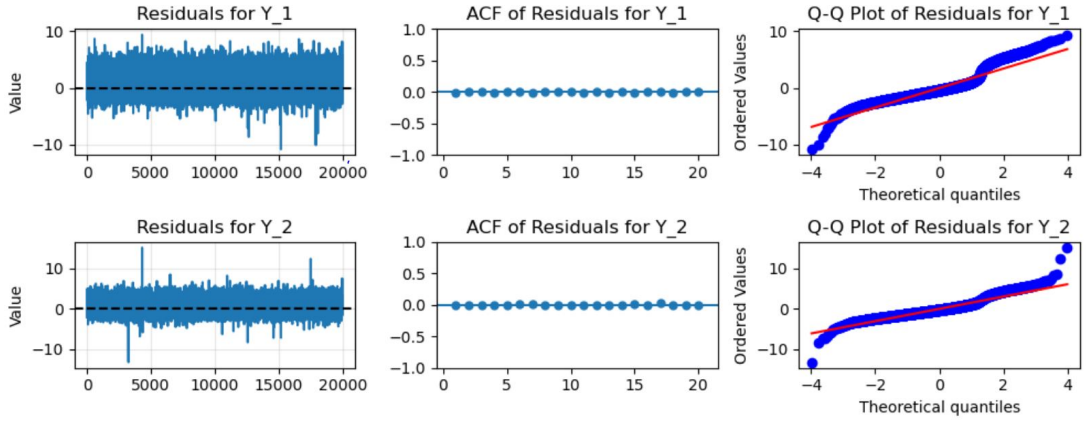
Under model diagnostics, the analysis of the residuals focused mainly on residual plots, normality of the residuals, and autocorrelation of residuals. Also, goodness-of-fits metrics were examined.

3.3.1 Residual analyses

In Figure 7, the ACF panels for both residuals show very low autocorrelation across lags, implying that the dynamic specification absorbs serial dependence in both cases. But the Extended



(A) Extended BVAR-ODC



(B) Standard BVAR

FIGURE 7. Residual Analysis of Extended and Standard BVAR Models

BVAR-ODC residuals show a bit of tighter clustering around zero and fewer extreme spikes, signaling better temporal whitening under the outlier-induced modeling. This is because robust likelihoods and shrinkage schemes reduce autocorrelation of residuals and improve model adequacy. [53, 56]. In the Extended BVAR-ODC model, the Q–Q plots fit a Student-t reference line quite closely across both variables, whereas the Q–Q plots of the Standard BVAR model deviate significantly in the tails under a normal-reference line. This means that unmodeled heavy tails inflate residual dispersion and degrade normal-based inference. Bayesian mixture models have

produce superior residual calibration and tail coverage in multivariate time series [57, 55]. Since residuals from the Standard BVAR are more dispersed and heavy-tailed, uncertainty measures like predictive intervals, error covariance estimates tends to be inflated, consistent with inflated variance estimates. Modeling heavy-tails and including outlier adjustments in the Extended BVAR-ODC yields residuals whose distribution is much closer to the model assumptions, enabling sharper and credible inference. This is consistent with the view of robust modeling in conditions of non-Gaussian shocks [58, 60]. The statistical tests in Table 3 supports the residual analyses of the two models. The poor residual behaviour in the Standard BVAR implies that

TABLE 3. Diagnostic Tests: Standard BVAR vs Extended BVAR-ODC

	Standard BVAR Model		**Extended BVAR-ODC Model**	
Variable of Interest	**p-value**	**Decision**	**p-value**	**Decision**
Ljung-Box Test (Lag=10)				
Y_1	0.289	No autocorrelation	0.077	No autocorrelation
Y_2	0.938	No autocorrelation	0.364	No autocorrelation
D'Agostino Normality Test				
Y_1	0.000	Not normal	0.076	Normal
Y_2	0.000	Not normal	0.368	Normal

Note: Ljung-Box Test H_0: No autocorrelation left after lag 10; D'Agostino Test H_0: Data is normally distributed

forecasts, impulse responses, and scenario simulations may inherit inflated uncertainty in shock estimation. The superior residual behaviour of the Extended BVAR-ODC indicates higher reliability in downstream inference [59, 55]. Therefore, the Extended BVAR-ODC with outlier adjustment provides more reliable inference, robust estimation, and better-calibrated predictive distributions, particularly in the presence of outliers.

3.3.2 Goodness-of-fit metrics

The goodness-of-fit are compared in Table 4. Across all metrics, the Extended BVAR-ODC performs consistently better, indicating improved fit and predictive accuracy. Consequently, the

Extended BVAR-ODC is superior across nearly all performance metrics. This indicates better fit, higher predictive accuracy, and stronger robustness against data irregularities than the Standard BVAR. The higher R^2 implies improved model fit. The lower MSE and MAE means more accurate and efficient model. This evidences align with the posterior predictive and residual analysis, which showed the Extended BVAR-ODC achieves improved predictive calibration, lower variance inflation, and better posterior stability. Hence, the Extended BVAR-ODC is statistically preferred, confirming its efficiency and robustness in capturing the underlying data-generating process.

TABLE 4. Goodness-of-Fit Measures and Information Criteria

Performance Measures	**Standard BVAR Model**	**Extended BVAR-ODC Model**
MSE	3.084	1.156
RMSE	1.756	1.075
MAE	1.225	0.815
MAD	0.851	0.678
R^2	0.691	0.884

In general, in the presence of outlier contamination, the standard BVAR posterior means for coefficients converge to skewed estimates, since the likelihood incorrectly assigns outlier effects to systematic dynamics [62], a finding corroborated by the simulation results in this study. The improved BVAR framework mitigates bias by explicitly modeling outliers, hence enhancing alignment with the theoretical prediction of asymptotically unbiased estimate when outliers are accurately recognized [63]. Theory suggests that unmodeled outliers inflate posterior uncertainty for VAR coefficients, as the model seeks to reconcile extreme observations with presumed Gaussian dynamics [64]. Empirical observations indicate that the 95% credible interval widths for the VAR coefficients are broader under the standard BVAR than the Extended BVAR-ODC.

The decrease in posterior uncertainty within the Extended BVAR-ODC indicates the effective differentiation of outlier influences from systematic dynamics, aligning with theoretical analysis [65].

4. Conclusions

The study has demonstrated that the Extended BVAR-ODC with outlier control through the hybrid DNN–DBSCAN technique provides a markedly superior fit compared to the Standard BVAR. Across all diagnostic levels, including parameter recovery, posterior trace convergence, and predictive calibration, the Extended BVAR-ODC consistently outperforms its variant. Particularly, posterior estimates of the intercepts, lagged coefficients, and covariance matrices exhibit higher stability and smaller posterior dispersion, while the p-values and t-values confirm stronger statistical significance and parameter credibility. The trace and posterior predictive plots reveal faster convergence and reduced sampling variability, indicating improved MCMC mixing. Moreover, predictive distribution diagnostics (mixture density, PIT, and CRPS) show the Extended BVAR-ODC achieves better alignment between observed and simulated data, ensuring both reliability and robustness. Finally, the residual analyses and goodness-of-fit confirm the Extended BVAR-ODC effectively mitigates the distortions induced by outliers, producing nearly white, homoscedastic residuals with heavier-tail adjustment through the Student-t component. Recommendations are made for the adoption of the Extended BVAR-ODC in modeling macroeconomic or financial systems subject to structural breaks, volatility shifts, or extreme events. Its outlier-induced specification ensures a more robust inference under contaminated data distributions. Also, the Standard BVAR implementations should always be complemented with posterior predictive and residual diagnostics before final inference to avoid biased conclusions due to the unmodeled tail behaviour. Lastly, the inclusion of model comparison metrics such as CRPS should become standard practice, ensuring transparent and reproducible model selection. Future work could provide comparative studies to assess the Extended BVAR-ODC against modern probabilistic deep learning models like Bayesian LSTMs or Temporal VAEs to evaluate trade-offs in interpretability, scalability, and uncertainty quantification.

Acknowledgments

The authors are grateful to the Pan African University Institute for Basic Sciences, Technology and Innovation (PAUSTI) for the support.

Conflict of Interest

The authors declare that there is no conflict of interests.

References

[1] S. Verstyuk, Modeling Multivariate Time Series in Economics: From Auto-Regressions to Recurrent Neural Networks, SSRN (2020). https://doi.org/10.2139/ssrn.3589337.

[2] J. Brownlee, Deep Learning for Time Series Forecasting Predict the Future with MLPs, CNNs and LSTMs in Python, Machine Learning Mastery, 2018. https://machinelearningmastery.com/deep-learning-for-time-series-forecasting.

[3] F. Khan, H. Iftikhar, I. Khan, P.C. Rodrigues, A.A. Alharbi, et al., A Hybrid Vector Autoregressive Model for Accurate Macroeconomic Forecasting: An Application to the U.S. Economy, Mathematics 13 (2025), 1706. https://doi.org/10.3390/math13111706.

[4] J.D. Hamilton, Time Series Analysis, Princeton University Press, 2020. https://doi.org/10.2307/j.ctv14jx6sm.

[5] J. Krampe, L. Margaritella, Factor Models with Sparse Vector Autoregressive Idiosyncratic Components, Oxf. Bull. Econ. Stat. 87 (2025), 837–849. https://doi.org/10.1111/obes.12664.

[6] D. Giannone, M. Lenza, G.E. Primiceri, Priors for the Long Run, J. Am. Stat. Assoc. 114 (2018), 565–580. https://doi.org/10.1080/01621459.2018.1483826.

[7] L.J. Christiano, Christopher A. Sims and Vector Autoregressions*, Scand. J. Econ. 114 (2012), 1082–1104. https://doi.org/10.1111/j.1467-9442.2012.01737.x.

[8] P. Esling, C. Agon, Time-Series Data Mining, ACM Comput. Surv. 45 (2012), 1–34. https://doi.org/10.1145/2379776.2379788.

[9] Z. Zhao, Z. Lai, H. Zhi, Y. Zou, Y. Jin, et al., Automated Workflow of EIS Data Validation and Quality Improvement Based on the Definition, Detection, and Removal of Outliers, Electrochimica Acta 461 (2023), 142661. https://doi.org/10.1016/j.electacta.2023.142661.

[10] Y. Yang, C. Fan, L. Chen, H. Xiong, IPMOD: An Efficient Outlier Detection Model for High-Dimensional Medical Data Streams, Expert Syst. Appl. 191 (2022), 116212. https://doi.org/10.1016/j.eswa.2021.116212.

[11] S. Sadik, Online Detection of Outliers for Data Streams, Phd Dissertation, University of Oklahoma, Norman, Oklahoma, (2013).

[12] M. Gupta, A. Mallya, S. Roy, J.H.D. Cho, J. Han, Local Learning for Mining Outlier Subgraphs from Network Datasets, in: Proceedings of the 2014 SIAM International Conference on Data Mining, Society for Industrial and Applied Mathematics, Philadelphia, PA, 2014, pp. 73–81. https://doi.org/10.1137/1.9781611973440.9.

[13] S. Ng, Modeling Macroeconomic Variations after Covid-19, NBER Working Paper No. 29060, National Bureau of Economic Research, Cambridge, Massachusetts, (2021).

[14] A. Carriero, T.E. Clark, M. Marcellino, E. Mertens, Addressing COVID-19 Outliers in BVARs with Stochastic Volatility, Rev. Econ. Stat. 106 (2024), 1403–1417. https://doi.org/10.1162/rest_a_01213.

[15] A. Carriero, J. Chan, T.E. Clark, M. Marcellino, Corrigendum to "Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors" [J. Econometrics 212 (1) (2019) 137–154], J. Econ. 227 (2022), 506–512. https://doi.org/10.1016/j.jeconom.2021.11.010.

[16] E. Bobeica, B. Hartwig, The COVID-19 Shock and Challenges for Time Series Models, ECB Working Paper Series, No. 2558, European Central Bank, (2021).

[17] A.G. Waititu, Nonparametric Changepoint Analysis for Bernoulli Random Variables Based on Neural Networks, Doctoral Thesis, Technische Universität Kaiserslautern (2008). https://kluedo.ub.rptu.de/frontdoor/index/index/docId/2032.

[18] G. Dorcas Wambui, The Power of the Pruned Exact Linear Time(PELT) Test in Multiple Changepoint Detection, Am. J. Theor. Appl. Stat. 4 (2015), 581–586. https://doi.org/10.11648/j.ajtas.20150406.30.

[19] B.K. Nkansah, Discordancy in Reduced Dimensions of Outliers in High-Dimensional Datasets: Application of an Updating Formula, Am. J. Theor. Appl. Stat. 2 (2013), 29–37. https://doi.org/10.11648/j.ajtas.20130202.14.

[20] T. Asamoah, A.G. Waititu, B.K. Nkansah, C. Omari, Outlier Detection in Multivariate Time Series: An Application of Hybrid Dnn-Dbscan Technique, Commun. Math. Biol. Neurosci. 2026 (2026), 9. https://doi.org/10.28919/cmbn/9659.

[21] A. Carriero, T.E. Clark, M. Marcellino, Common Drifting Volatility in Large Bayesian VARs, J. Bus. Econ. Stat. 34 (2016), 375–390. https://doi.org/10.1080/07350015.2015.1040116.

[22] E. Aamari, C. Berenfeld, C. Levrard, Optimal Reach Estimation and Metric Learning, Ann. Stat. 51 (2023), 1148–1180. https://doi.org/10.1214/23-AOS2281.

[23] F. Huber, G. Kastner, M. Feldkircher, Should I Stay or Should I Go? A Latent Threshold Approach to Large-scale Mixture Innovation Models, J. Appl. Econ. 34 (2019), 621–640. https://doi.org/10.1002/jae.2680.

[24] C. Ramos-Carreño, J.L. Torrecilla, M. Carbajo-Berrocal, P. Marcos, A. Suárez, scikit-fda: A Python Package for Functional Data Analysis, J. Stat. Softw. 109 (2024), 1–37. https://doi.org/10.18637/jss.v109.i02.

[25] J.L. Cross, C. Hou, A. Poon, Macroeconomic Forecasting with Large Bayesian VARs: Global-Local Priors and the Illusion of Sparsity, Int. J. Forecast. 36 (2020), 899–915. https://doi.org/10.1016/j.ijforecast.2019.10.002.

[26] C. Francq, J. Zakoian, GARCH Models, Wiley, 2019. https://doi.org/10.1002/9781119313472.

[27] C.C. Aggarwal, Outlier Analysis, Springer, 2017. https://doi.org/10.1007/978-3-319-47578-3.

[28] G. Koop, D. Korobilis, Bayesian Multivariate Time Series Methods for Empirical Macroeconomics, Found. Trends Econ. 3 (2010), 267–358. https://doi.org/10.1561/0800000013.

[29] Q. Chen, Y. Hong, H. Li, Time-Varying Forecast Combination for Factor-Augmented Regressions with Smooth Structural Changes, J. Econ. 240 (2024), 105693. https://doi.org/10.1016/j.jeconom.2024.105693.

[30] C.A.W. Glas, T.D. Jorgensen, D.T. Hove, Reducing Attenuation Bias in Regression Analyses Involving Rating Scale Data via Psychometric Modeling, Psychometrika 89 (2024), 42–63. https://doi.org/10.1007/s11336-024-09967-4.

[31] K.L. Lange, R.J.A. Little, J.M.G. Taylor, Robust Statistical Modeling Using the t Distribution, J. Am. Stat. Assoc. 84 (1989), 881–896. https://doi.org/10.1080/01621459.1989.10478852.

[32] M. Marcellino, J.H. Stock, M.W. Watson, Macroeconomic Forecasting in the Euro Area: Country Specific Versus Area-Wide Information, Eur. Econ. Rev. 47 (2003), 1–18. https://doi.org/10.1016/S0014-2921(02)00206-4.

[33] E. Schubert, J. Sander, M. Ester, H.P. Kriegel, X. Xu, DBSCAN Revisited, Revisited: Why and How You Should (Still) Use DBSCAN, ACM Trans. Database Syst. 42 (2017), 1–21. https://doi.org/10.1145/3068335.

[34] L. Bontemps, V.L. Cao, J. McDermott, N.A. Le-Khac, Collective Anomaly Detection Based on Long Short-Term Memory Recurrent Neural Networks, in: Lecture Notes in Computer Science, Springer, Cham, 2016: pp. 141–152. https://doi.org/10.1007/978-3-319-48057-2_9.

[35] A. Carriero, T.E. Clark, M. Marcellino, Nowcasting Tail Risk to Economic Activity at a Weekly Frequency, J. Appl. Econ. 37 (2022), 843–866. https://doi.org/10.1002/jae.2903.

[36] N. Kuschnig, L. Vashold, BVAR: Bayesian Vector Autoregressions with Hierarchical Prior Selection in R, J. Stat. Softw. 100 (2021), 1–27. https://doi.org/10.18637/jss.v100.i14.

[37] J.C. Chan, Minnesota-Type Adaptive Hierarchical Priors for Large Bayesian VARs, Int. J. Forecast. 37 (2021), 1212–1226. https://doi.org/10.1016/j.ijforecast.2021.01.002.

[38] D. Gefang, G. Koop, A. Poon, Computationally Efficient Inference in Large Bayesian Mixed Frequency VARs, Econ. Lett. 191 (2020), 109120. https://doi.org/10.1016/j.econlet.2020.109120.

[39] D. Korobilis, Quantile Regression Forecasts of Inflation Under Model Uncertainty, Int. J. Forecast. 33 (2017), 11–20. https://doi.org/10.1016/j.ijforecast.2016.07.005.

[40] S. Ankargren, M. Unosson, Y. Yang, A Flexible Mixed-Frequency Vector Autoregression with a Steady-State Prior, J. Time Ser. Econ. 12 (2020), 20180034. https://doi.org/10.1515/jtse-2018-0034.

[41] F.O. Lopez, A Bayesian Approach to Parameter Estimation in Simplex Regression Model: A Comparison with Beta Regression, Rev. Colomb. Estad. 36(2013), 1–21.

[42] J.C. Chan, X. Yu, W. Zhang, Bayesian Model Comparison for Large Bayesian VARs After the COVID-19 Pandemic, J. Econ. (2025), 106072. https://doi.org/10.1016/j.jeconom.2025.106072.

[43] M. Bańbura, D. Giannone, M. Modugno, L. Reichlin, Now-Casting and the Real-Time Data Flow, in: Handbook of Economic Forecasting, Elsevier, 2013: pp. 195–237. https://doi.org/10.1016/B978-0-444-53683-9.00004-9.

[44] F. Huber, G. Koop, L. Onorante, M. Pfarrhofer, J. Schreiner, Nowcasting in a Pandemic Using Non-Parametric Mixed Frequency VARs, J. Econ. 232 (2023), 52–69. https://doi.org/10.1016/j.jeconom.2020.11.006.

[45] A. Carriero, T.E. Clark, M. Marcellino, Large Bayesian Vector Autoregressions with Stochastic Volatility and Non-Conjugate Priors, J. Econ. 212 (2019), 137–154. https://doi.org/10.1016/j.jeconom.2019.04.024.

[46] S. Frühwirth-Schnatter, G. Malsiner-Walli, From Here to Infinity: Sparse Finite Versus Dirichlet Process Mixtures in Model-Based Clustering, Adv. Data Anal. Classif. 13 (2018), 33–64. https://doi.org/10.1007/s11634-018-0329-y.

[47] M. Bańbura, D. Giannone, L. Reichlin, Large Bayesian Vector Auto Regressions, J. Appl. Econ. 25 (2010), 71–92. https://doi.org/10.1002/jae.1137.

[48] W. Liu, H. Fan, M. Xia, Tree-Based Heterogeneous Cascade Ensemble Model for Credit Scoring, Int. J. Forecast. 39 (2023), 1593–1614. https://doi.org/10.1016/j.ijforecast.2022.07.007.

[49] A. Vehtari, A. Gelman, J. Gabry, Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC, Stat. Comput. 27 (2016), 1413–1432. https://doi.org/10.1007/s11222-016-9696-4.

[50] J. Gabry, D. Simpson, A. Vehtari, M. Betancourt, A. Gelman, Visualization in Bayesian Workflow, J. R. Stat. Soc. Ser.: Stat. Soc. 182 (2019), 389–402. https://doi.org/10.1111/rssa.12378.

[51] M. Billio, R. Casarin, L. Rossini, Bayesian Nonparametric Sparse VAR Models, J. Econ. 212 (2019), 97–115. https://doi.org/10.1016/j.jeconom.2019.04.022.

[52] J. Doornik, D.F. Hendry, GiveWin: Version 2: An Interface to Empirical Modelling, Timberlake Consultants Press, 1999.

[53] J.C.C. Chan, Asymmetric Conjugate Priors for Large Bayesian VARs, Quant. Econ. 13 (2022), 1145–1169. https://doi.org/10.3982/QE1381.

[54] R.J.G.B. Campello, D. Moulavi, A. Zimek, J. Sander, Hierarchical Density Estimates for Data Clustering, Visualization, and Outlier Detection, ACM Trans. Knowl. Discov. Data 10 (2015), 1–51. https://doi.org/10.1145/2733381.

[55] M. Mayrhofer, U. Radojičić, P. Filzmoser, Robust Covariance Estimation and Explainable Outlier Detection for Matrix-Valued Data, arXiv:2403.03975, 2024. https://doi.org/10.48550/arXiv.2403.03975.

[56] J. Prüser, B. Blagov, Improving Inference and Forecasting in VAR Models Using Cross-Sectional Information, Econ. Model. 160 (2026), 107618. https://doi.org/10.1016/j.econmod.2026.107618.

[57] P. Gagnon, Y. Hayashi, Theoretical Properties of Bayesian Student-t Linear Regression, arXiv:2204.02299, 2022. https://doi.org/10.48550/arXiv.2204.02299.

[58] B. Grün, G. Malsiner-Walli, Bayesian Finite Mixture Models, arXiv:2407.05470, 2024. https://doi.org/10.48550/arXiv.2407.05470.

[59] C. Retiti Diop Emane, S. Song, H. Lee, D. Choi, J. Lim, et al., Anomaly Detection Based on GCNs and DBSCAN in a Large-Scale Graph, Electronics 13 (2024), 2625. https://doi.org/10.3390/electronics13132625.

[60] M.W. McCracken, M.T. Owyang, T. Sekhposyan, Real-Time Forecasting and Scenario Analysis Using a Large Mixed-Frequency Bayesian VAR, Int. J. Central Bank. 17 (2021), 327–367.

[61] Z. Jiao, P. Hu, H. Xu, Q. Wang, Machine Learning and Deep Learning in Chemical Health and Safety: A Systematic Review of Techniques and Applications, ACS Chem. Health Saf. 27 (2020), 316–334. https://doi.org/10.1021/acs.chas.0c00075.

[62] M. Riani, A. Corbellini, A.C. Atkinson, The Use of Prior Information in Very Robust Regression for Fraud Detection, Int. Stat. Rev. 86 (2018), 205–218. https://doi.org/10.1111/insr.12247.

[63] A. Desgagné, P. Gagnon, Bayesian Robustness to Outliers in Linear Regression and Ratio Estimation, Braz. J. Probab. Stat. 33 (2019), 205–221. https://doi.org/10.1214/17-bjps385.

[64] D. Creal, B. Schwaab, S.J. Koopman, A. Lucas, Observation-Driven Mixed-Measurement Dynamic Factor Models with an Application to Credit Risk, Rev. Econ. Stat. 96 (2014), 898–915. https://doi.org/10.1162/REST_a_00393.

[65] S. Das, M.G. Genton, Y.M. Alshehri, G.L. Stenchikov, A Cyclostationary Model for Temporal Forecasting and Simulation of Solar Global Horizontal Irradiance, Environmetrics 32 (2021), e2700. https://doi.org/10.1002/env.2700.