

# Short Term Forecasting Performances of Classical VAR and Sims-Zha Bayesian VAR Models for Time Series with Collinear Variables and Correlated Error Terms

M. O. Adenomon\*, V. A. Michael, O. P. Evans

Department of Mathematics & Statistics, The Federal Polytechnic, Bida, Nigeria

Email: \*admonsagie@gmail.com, vicronz@yahoo.com, patevansjj@gmail.com

Received 11 August 2015; accepted 15 December 2015; published 18 December 2015

Copyright © 2015 by authors and Scientific Research Publishing Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

---

## Abstract

Forecasts can either be short term, medium term or long term. In this work, we consider short term forecast because of the problem of limited data or time series data that often encounter in time series analysis. This simulation study considers the performances of the classical VAR and Sims-Zha Bayesian VAR for short term series at different levels of collinearity and correlated error terms. The results from 10,000 iteration reveal that the BVAR models are excellent for time series length of  $T = 8$  for all levels of collinearity while the classical VAR is effective for time series length of  $T = 16$  for all collinearity levels except when  $\rho = -0.9$  and  $\rho = -0.95$ . Therefore, we recommend that for effective short term forecasting, the time series length, forecasting horizon and the collinearity level should be considered.

## Keywords

Short term, Forecasting, Vector Autoregressive (VAR), Bayesian VAR (BVAR), Sims-Zha Prior, Collinearity, Error Terms

---

## 1. Introduction

Forecasts from classical Vector Autoregression (VAR) models and the Bayesian VAR had gained great popularity in the 1980s ([1] [2]). In recent times, the classical VAR models tend to over fit the data and overestimate the magnitude of the coefficients of distant lags of variables as a result of sampling error [3]. But the Bayesian VAR

\*Corresponding author.

is designed to improve macroeconomic forecast and to solve many problems associated with the classical VAR. The many advantages of the BVAR make it more useful in forecasting short term macroeconomic series. As observed by Bischoff *et al.*, [2] that a small VAR, Bayesian or otherwise, can be useful for the purpose of forecast only.

Forecast can either be short term, medium term or long term. In this present work, our focus is on short term forecast because of the problem of limited data or time series that may be encountered in time series analysis.

Short term forecasting is very useful for decision making in many fields of life. [4] uses the Bayesian VAR to model and forecast the intraday electricity load in the short term. [5] considers short term load forecasting methods based on European data, while [6] [7] and [8] use different statistical models to forecast GDP in their respective countries. [9] uses factor models for short term forecasting of the Japanese economy; [10] considers factors affecting short term load forecasting and, [11] in their review considers short term weather forecasting from the physic and mathematics point of view. For other examples on short term forecasting see [12] [13] and [14].

This present work is motivated by the work of Johnson, [15] who studies the effect of correlation and identification status on methods of estimating parameters of system of simultaneous equations using Monte Carlo approach. Also the work of Bischoff, Balay and Kang, [2] give rise to this present work. Bischoff, Balay and Kang suggest that if one only wants to forecast (*i.e.* not do “structural analysis”) a small VAR, Bayesian, otherwise, can be useful. Lastly this work is also motivated by the recent work by Dormann *et al.*, [16] who observe in their simulation studies on GLM that correlation coefficient of  $|r| < 0.7$  is an appropriate indicator when collinearity begins to severely distort model estimation and subsequent prediction.

Therefore, the aim of this paper is to compare the performances of the classical VAR and Bayesian VAR for time series with collinear variables and correlated error terms in the short term.

## 2. Review of Related Literatures

The name “multicollinearity” was first introduced by Ragnar Frisch. In his original formulation the economic variables are supposed to be composed of two parts, a systematic or “true” and an “error” component. This problem which arise when some or all the variables in the regression equation are highly intercorrelated and it becomes also impossible to separate their influences and obtain the corresponding estimates of the regression coefficient [17]. Multicollinearity is a term that refers to correlation among independent variables in a multiple regression model; it is usually invoked when some correlations are “large” but an actual magnitude test is not well defined [18].

[19] proposed a statistic that measure the ‘distance’ of a cross-product matrix from the diagonal matrix obtained by zeroing its off-diagonal elements and they found it useful in detecting near multicollinearity regression problems. It can also distinguished between apparent and real multicollinearity with positive probability.

[20] worked out the statistical implications of the orthogonalization procedure in the general linear model. His work demonstrated that orthogonalization can worsen collinearity if measured by its effect on estimated variances.

[21] reported that parameter estimates may be poorly determined in-sample due to the sheer number of variables, perhaps worsened by the high degree of collinearity manifested in the levels of integrated data.

[22] investigated the non-uniqueness of collinearity using the static regression model and reported that any collinearity in the explanatory variables is irrelevant to forecasting so long as the marginal process remains constant.

[23] noted that since non-experimental data in general and economics data in particular, are often highly correlated, then from Bayesian viewpoint model specification is closely related to the problem of multicollinearity. In their approach they compared the predictive densities for an equation with and without the set of variables in question in order to gauge that the set may be safely omitted if the omission has little or no effect on the predictive densities. They concluded that examination of changes in predictive means and of Generalized Variance Ratio (GVR) is a useful method of investigation model specification.

[24] revealed through their Monte Carlo simulation experiment that multicollinearity can cause problems in theory testing (Type II errors) under certain conditions which includes: when multicollinearity is extreme; when multicollinearity is between 0.6 and 0.8, and when multicollinearity is between 0.4 and 0.5.

[25] studied graphical views of suppressor variables and multicollinearity in multiple linear regression.

[26] investigated the Type II error rate of the OLS estimators at different levels of multicollinearity and sample sizes through Monte-Carlo studies. Their work revealed that increasing the sample size reduces the type II error rate of the OLS estimator at all levels of multicollinearity.

[27] considered the available options available to researchers when one or more assumptions of an ordinary Least Squares (OLS) regression model are violated. Ayyangar paid particular attention on the problems of skewness, multicollinearity and heteroskedasticity and autocorrelated error terms on OLS models using SAS with illustration to health care cost data.

[15] studied the effects of correlation and identification status on methods of estimating parameters of system of simultaneous equations model using Monte Carlo approach. The Monte Carlo approach for the performances of the estimating methods at different levels of correlation, sample sizes and identification status were reported.

[28] considered the various performances of Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC) and Mallow  $C_p$  statistic in the presence of a multicollinear regression using simulated data from SAS programme. Their work revealed that the performances of AIC and BIC in choosing the correct model among collinear variables are better when compared with the performances of Mallow's  $C_p$ .

[29] studied the effects of multicollinearity and the sensitivity of the estimation methods in simultaneous equation model for different levels of different levels of multicollinearity. He considered Ordinary Least Squares (OLS), 2 Stage Least Squares (2SLS) and 3 Stage Least Squares (3SLS) methods of estimation. His result revealed preference of 3SLS over 2SLS and OLS.

[30] noted that multicollinearity problems are almost always present in time-series data generated by natural experiments. He also noted that multicollinearity becomes "harmful" when there is an  $R^2$  in the predictor matrix that is of the same order of magnitude as the  $R^2$  of the overall model.

[31] studied the effect of multicollinearity on some estimators (Ordinary Least Squares, Cochran-Orcut (GLS2), Maximum Likelihood Estimator (MLE), Multivariate Regression, Full Information Maximum Likelihood, Seemingly Unrelated Regression (SUR) and Three Stage Least Squares (3SLS)). Results showed that multivariate regression, FIML, SUR and 3SLS estimators are preferred at all levels of sample size.

[32] studied the performances of Bayesian Linear regression, Ridge Regression and OLS methods for modeling collinear data.

[16] who observed in their simulation studies on GLM that correlation coefficient of  $|r| < 0.7$  was an appropriate indicator for when collinearity begins to severely distort model estimation and subsequent prediction.

[33] compared the forecasting performances of the Reduced form Vector Autoregression (VAR) and Sims-Zha Bayesian VAR (BVAR) in a situation where the endogenous variables are collinear at different levels and at different short terms time series lengths assuming harmonic decay.

Sources of multicollinearity includes: the data collection method employed; constraints on the model or in the population being sampled; model specification; and overdetermined model. Multicollinearity especially in time series data may occur if the regressors included in the model share a common trend, that is, they all increase or decrease over time.

[34] identified some consequences of multicollinearity. They include:

- 1) Although BLUE, the OLS estimators have large variances and covariances making precise estimation difficult.
- 2) Because of consequence 1, the confidence interval tends to be much wider, leading to the acceptance of the (zero null hypothesis) and the  $t$ -ratio of one or more coefficients tends to be statistically insignificant.
- 3) Although the  $t$ -ratio of one or more coefficients is statistically insignificant,  $R^2$ , the overall measure of goodness-of-fit can be very high.
- 4) The OLS estimators and their standard errors can be sensitive to small change in the data.

### 3. Model Description

#### 3.1. Vector Autoregression (VAR) Model

VAR methodology superficially resembles simultaneous equation modeling in that we consider several endogenous variables together. But each endogenous variable is explained by its lagged values and the lagged values of all other endogenous variables in the model; usually, there are no exogenous variables in the model [34].

Given a set of  $k$  time series variables,  $y_t = [y_{1t}, \dots, y_{kt}]'$ , VAR models of the form

$$y_t = A_1 y_{t-1} + \dots + A_p y_{t-p} + u_t \tag{1}$$

provide a fairly general framework for the Data General Process (DGP) of the series. More precisely this model is called a VAR process of order  $p$  or VAR( $p$ ) process. Here  $u_t = [u_{1t}, \dots, u_{kt}]'$  is a zero mean independent white noise process with non singular time invariant covariance matrix  $\Sigma_u$  and the  $A_i$  are  $(k \times k)$  coefficient matrices. The process is easy to use for forecasting purpose though it is not easy to determine the exact relations between the variables represented by the VAR model in Equation (1) above [35]. Also, polynomial trends or seasonal dummies can be included in the model.

The process is stable if

$$\det(I_k - A_1 z - \dots - A_p z^p) \neq 0 \text{ for } |z| \leq 1 \tag{2}$$

In that case it generates stationary time series with time invariant means and variance covariance structure. Therefore To estimate the VAR model, one can write a VAR( $p$ ) with a concise matrix notation as

$$Y = BZ + U$$

where  $Y = [y_1, \dots, y_T]$ ,  $Z_{t-1} = \begin{bmatrix} y_{t-1} \\ \vdots \\ y_{t-p} \end{bmatrix}$ ,  $Z = [Z_0, \dots, Z_{T-1}]$  (3)

Then the Multivariate Least Squares (MLS) for  $B$  yields

$$\hat{B} = (ZZ')^{-1} Z'Y \tag{4}$$

It can be written alternatively as

$$Vec(\hat{B}) = ((ZZ')^{-1} Z \otimes I_k) Vec(Y) \tag{5}$$

where  $\otimes$  denotes the Kronecker product and  $Vec$  the vectorization of the matrix  $Y$ . This estimator is consistent and asymptotically efficient. It furthermore equals the conditional Maximum Likelihood Estimator (MLE) [36].

As the explanatory variables are the same in each equation, the multivariate least squares is equivalent to the Ordinary Least Squares (OLS) estimator applied to each equation separately, as was shown by [37].

In the standard case, the MLE estimator of the covariance matrix differs from the OLS estimator.

$$\text{MLE estimator } \hat{\Sigma} = \frac{1}{T} \sum_{t=1}^T \hat{\ell}_t \hat{\ell}_t' \tag{6}$$

OLS estimator for a model with a constant,  $k$  variables and  $p$  lags, in a matrix notation, gives

$$\hat{\Sigma} = \frac{1}{T - kp - 1} (Y - \hat{B}Z)(Y - \hat{B}Z)' \tag{7}$$

Therefore, the covariance matrix of the parameters can be estimated as

$$C\hat{O}v = (Vec(\hat{B})) = (ZZ')^{-1} \otimes \hat{\Sigma} \tag{8}$$

### 3.2. Bayesian Vector Autoregression with Sims-Zha Prior

In recent times, the BVAR model of [38] has gained popularity both in economic time series and political analysis. The Sims-Zha BVAR allows for a more general specification and can produce a tractable multivariate normal posterior distribution. Again, the Sims-Zha BVAR estimates the parameters for the full system in a multivariate regression [3].

Given the reduced form model

$$y_t = c + y_{t-1}B_1 + \dots + y_{t-p}B_p + u_t$$

where  $c = dA_0^{-1}$ ,  $B_l = -A_l A_0^{-1}$ ,  $l = 1, 2, \dots, p$ ,  $u_t = \varepsilon_t A_0^{-1}$  and  $\Sigma = A_0^{-1} A_0^{-1}$

The matrix representation of the reduced form is given as

$$Y = \begin{matrix} X \\ T \times m \end{matrix} \begin{matrix} \beta \\ T \times (mp+1) \end{matrix} + \begin{matrix} U \\ (mp+1) \times m \\ T \times m \end{matrix}, \quad U \sim MVN(0, \Sigma)$$

We can then construct a reduced form Bayesian SUR with the Sims-Zha prior as follows. The prior means for the reduced form coefficients are that  $B_1 = I$  and  $B_2, \dots, B_p = 0$ . We assume that the prior has a conditional structure that is multivariate Normal-inverse Wishart distribution for the parameters in the model. To estimate the coefficients for the system of the reduced form model with the following estimators

$$\hat{\beta} = (\Psi^{-1} + X'X)^{-1} (\Psi^{-1}\bar{\beta} + X'Y)$$

$$\hat{\Sigma} = T^{-1} (Y'Y - \hat{\beta}'(X'X + \Psi^{-1})\hat{\beta} + \bar{\beta}'\Psi^{-1}\bar{\beta} + \bar{S})$$

where the Normal-inverse Wishart prior for the coefficients is

$$\beta/\Sigma \sim N(\bar{\beta}, \Psi) \text{ and } \Sigma \sim IW(\bar{S}, \nu)$$

This representation translates the prior proposed by Sims-Zha form from the structural model to the reduced form ([3] [39] and [38] [40]).

The summary of the Sims-Zha prior is given in **Table 1**.

### 3.3. Simulation Procedure

The simulated data will be generated for time series lengths of 8 and 16. The choice of the length chosen is to be able to study the models in the short run [41]. This study also considered ten (10) multicollinearity levels as  $\rho = (0.8, -0.8, 0.85, -0.85, 0.9, -0.9, 0.95, -0.95, 0.99, -0.99)$ .

The simulation procedure is given in the following steps.

**Step1:** we generated a VAR (2) process that obeys the following form

$$\begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_t = \begin{bmatrix} 5.0 \\ 10.0 \end{bmatrix} + \begin{bmatrix} 0.5 & 0.2 \\ -0.2 & -0.5 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-1} + \begin{bmatrix} -0.3 & -0.7 \\ -0.1 & 0.3 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \end{bmatrix}_{t-2} + \begin{bmatrix} u_1 \\ u_2 \end{bmatrix}_t$$

Our choice for this form model is to obtain a stable process and a VAR process that is not affected by overparameterization [8].

**Step2:** let the desired correlation matrix be  $R = \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$  then the Choleski factor  $P$  is  $P = \begin{bmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{bmatrix}$

and then the simulated data in **Step 1** is pre-multiplied by the Choleski factor so that the simulated data is scaled to have the desired correlation level [41].

**Step 3:** the VAR and BVAR models of lag length of 2 was be used for modeling and forecasting simulated data to obtain the RMSE and MAE.

**Step 4:** **Step 1 to Step 3** was repeated for 10,000 times, and the averages of the criteria were used to access the preferred model. A sample of simulated data is presented in **Table 2** below.

**Table 1.** Hyperparameters of Sims-Zha reference prior.

Parameter	Range	Interpretation	Range	Interpretation
$\lambda_0$			[0, 1]	Overall scale of the error covariance matrix
$\lambda_1$			>0	Standard deviation around $A_1$ (persistence)
$\lambda_2$			=1	Weight of own lag versus other lags
$\lambda_3$			>0	Lag decay
$\lambda_4$			$\geq 0$	standard deviation of intercept
$\lambda_5$			$\geq 0$	Scale of standard deviation of exogenous variable coefficients
$\mu_5$			$\geq 0$	Sum of coefficients/Cointegration (long-term trends)
$\mu_6$			$\geq 0$	Initial observations/dummy observation (impacts of initial conditions)
$\nu$			>0	Prior degrees of freedom

Source: Brandt and Freeman, [3].

**Table 2.** Sample of Simulated data for  $T = 16$ .

T	$\rho = -0.99$		Residuals	
	$y_1$	$y_2$	[U, 1]	[U, 2]
1	6.0853378	-4.699854321		
2	8.3361587	-7.662052368	-0.7440477	0.7748210
3	-1.4119188	2.549041262	0.2270471	-0.3013146
4	1.6662033	-0.878217866	0.5898282	-0.6190964
5	1.9049119	-0.455346148	0.6080660	-0.6421680
6	4.5498919	-3.60183345	1.4032863	-1.5356158
7	1.836577	-0.801168974	0.2527919	-0.1563110
8	1.7007428	-0.544318103	0.1370333	-0.1499999
9	2.2602923	-1.158178434	0.4050407	-0.3836562
10	2.0551198	-0.897890179	0.2085433	-0.1174189
11	2.0660043	-0.891296203	0.7917755	-0.7295352
12	2.5693203	-1.430130241	-1.4718452	1.3755778
13	0.1955522	0.800465005	-0.2213798	0.2439514
14	0.6950544	0.494571004	-1.4442621	1.4103745
15	1.1474605	0.006271441	-0.7418776	0.8303912
16	1.043101	0.203232356		
Estimated correlation $r = -0.9970061$		Estimated correlation of the residual $r = -0.996$		

### 3.4. Model Specification

The time series were generated data using a VAR model with lag 2. The choice here is to obtain a bivariate time series with the true lag length. While the VAR and BVAR models of lag length of 2 will be used for modeling and forecasting purpose.

For the BVAR model with Sims-Zha prior, we will consider the following range of values for the hyperparameters given below and the Normal-Inverse Wishart prior.

We consider two tight priors and two loose priors as follows:

The Tight priors are as follows

$$\text{BVAR1} = (\lambda_0 = 0.6, \lambda_1 = 0.1, \lambda_3 = 2, \lambda_4 = 0.1, \lambda_5 = 0.07, \mu_5 = \mu_6 = 5)$$

$$\text{BVAR2} = (\lambda_0 = 0.8, \lambda_1 = 0.1, \lambda_3 = 2, \lambda_4 = 0.1, \lambda_5 = 0.07, \mu_5 = \mu_6 = 5)$$

The Loose priors are as follows

$$\text{BVAR3} = (\lambda_0 = 0.6, \lambda_1 = 0.15, \lambda_3 = 2, \lambda_4 = 0.15, \lambda_5 = 0.07, \mu_5 = \mu_6 = 2)$$

$$\text{BVAR4} = (\lambda_0 = 0.8, \lambda_1 = 0.15, \lambda_3 = 2, \lambda_4 = 0.15, \lambda_5 = 0.07, \mu_5 = \mu_6 = 2)$$

where  $n\mu$  is prior degrees of freedom given as  $m + 1$  where  $m$  is the number of variables in the multiple time series data. In work  $n\mu$  is 3 (that is two (2) time series variables plus 1(one)).

The choice of Normal-inverse Wishart prior for the BVAR models follow the work of [42] that Normal Wishart prior tends to performed better when compared to other priors. In addition [38] proposed Normal-Inverse Wishart prior because of its suitability for large systems while [43] reported that the most advantage of Wishart distribution is that it guaranteed to produce positive definite draws. Our choice of the overall tightness  $\lambda_0 = 0.6$  and  $0.8$  is in line with work of [44]. In this work we assumed that the bivariate time series follows a quadratic decay. The Quadratic Decay (QD) model has many attractive theoretical properties that is why it is been applied to many fields of endeavour ([45] [46] and [47]).

The following are the criteria for forecast assessments used:

- 1) Mean Absolute Error (MAE) has a formular  $\text{MAE}_j = \frac{\sum_{i=1}^n |e_i|}{n}$ . This criterion measures deviation from the

series in absolute terms, and measures how much the forecast is biased. This measure is one of the most common ones used for analyzing the quality of different forecasts.

2) The Root Mean Square Error (RMSE) is given as  $RMSE_j = \sqrt{\frac{\sum_i^n (y_i - y^f)^2}{n}}$  where  $y_i$  is the time series data and  $y^f$  is the forecast value of  $y$  [8].

For the two measures above, the smaller the value, the better the fit of the model [48].

In this simulation study,  $RMSE = \frac{\sum_j^N RMSE_j}{N}$  and  $MAE = \frac{\sum_j^N MAE_j}{N}$  where  $N = 10,000$ . Therefore, the model with the minimum RMSE and MAE result as the preferred model.

### 3.5. Statistical Packages (R)

In this study three procedures in the R package will be used. They are: Dynamic System Estimation (DSE) [49]; the vars [50]; and the MSBVAR [51].

## 4. Results and Discussion

The results from the analysis are presented in **Tables 3-5**. The values of the RMSE and MAE are presented in **Table 3**. Generally the values of the RMSE and MAE decrease as a result of increase in the time series length from 8 to 16. Also the values of the RMSE and MAE increase as the collinearity values increases.

**Table 3.** Forecasting performances of the models for different time series length and forecasting horizons at different levels of collinearity.

Collinearity Levels ( $\rho$ )	Models	T = 8		T = 16	
		RMSE	MAE	RMSE	MAE
-0.99	VAR(2)	N/A	N/A	2.822276	1.865384
	BVAR1	3.480238	2.459911	2.924542	2.070996
	BVAR2	3.475351	2.461614	2.915648	2.064497
	BVAR3	3.487892	2.47295	2.8539	1.969234
	BVAR4	3.520224	2.502497	2.846255	1.958829
-0.95	VAR(2)	N/A	N/A	2.92741	1.923175
	BVAR1	3.513841	2.461495	2.955763	2.077187
	BVAR2	3.524945	2.48011	2.93884	2.058831
	BVAR3	3.463551	2.424384	2.808864	1.903786
	BVAR4	3.477056	2.428536	2.779634	1.878388
-0.9	VAR(2)	N/A	N/A	2.839813	1.892972
	BVAR1	3.527144	2.461289	2.957319	2.080047
	BVAR2	3.510038	2.455105	2.946401	2.06848
	BVAR3	3.453288	2.416771	2.80427	1.932373
	BVAR4	3.423209	2.39757	2.785012	1.918003
-0.85	VAR(2)	N/A	N/A	2.803335	1.871187
	BVAR1	3.503068	2.448057	2.959128	2.098831
	BVAR2	3.512622	2.462979	2.967651	2.114981
	BVAR3	3.422623	2.410443	2.846845	2.006635
	BVAR4	3.406495	2.409279	2.831411	2.009857

Continued

	VAR(2)	N/A	N/A	2.794558	1.871746
	BVAR1	3.499985	2.469094	3.003398	2.172383
-0.8	BVAR2	3.485945	2.4724	3.023571	2.216076
	BVAR3	3.421835	2.443648	2.917568	2.109152
	BVAR4	3.407584	2.454034	2.893364	2.10613
	VAR(2)	N/A	N/A	2.319124	1.560363
	BVAR1	3.048499	2.207239	2.551886	1.87176
0.8	BVAR2	3.045309	2.218768	2.575611	1.904332
	BVAR3	3.291741	2.559531	2.927194	2.326211
	BVAR4	3.345142	2.619884	2.879254	2.269554
	VAR(2)	N/A	N/A	2.383968	1.599955
	BVAR1	3.088592	2.237249	2.593269	1.898944
0.85	BVAR2	3.115247	2.269976	2.607954	1.921699
	BVAR3	3.331697	2.580985	2.953197	2.340492
	BVAR4	3.391991	2.652839	2.908163	2.285598
	VAR(2)	N/A	N/A	2.446215	1.638628
	BVAR1	3.159379	2.282306	2.653275	1.934046
0.9	BVAR2	3.159299	2.287451	2.657339	1.953146
	BVAR3	3.372884	2.588348	2.970293	2.33445
	BVAR4	3.435369	2.66265	2.955572	2.308289
	VAR(2)	N/A	N/A	2.526335	1.683815
	BVAR1	3.239628	2.325308	2.715885	1.968942
0.95	BVAR2	3.255642	2.353053	2.726984	1.993325
	BVAR3	3.423552	2.582054	2.986378	2.305256
	BVAR4	3.480364	2.651436	2.97899	2.290409
	VAR(2)	N/A	N/A	2.613225	1.739059
	BVAR1	3.346564	2.381411	2.796383	2.012807
0.99	BVAR2	3.35575	2.405958	2.820216	2.040734
	BVAR3	3.457429	2.546992	2.954774	2.197399
	BVAR4	3.531113	2.624224	2.965086	2.200766

In **Table 4** below, we present the ranks of the performances of the model. Rank of 1 is assigned to the model with the smallest values of RMSE and MAE while rank of 5 is assigned to the model with the largest values of the RMSE and MAE.

In **Table 5** below, we present the preferred model for the different time series length and forecasting horizons at the different collinearity levels. It was observed that from  $\rho = -0.95$  to  $\rho = -0.8$  for  $T = 8$ , the BVAR model with loose prior are preferred while for  $\rho = -0.99$ , the BVAR model with tight prior is preferred. For the positive collinearity from  $\rho = 0.8$  to  $\rho = 0.99$  for  $T = 8$ , the BVAR model with tight prior are preferred. In general for  $T = 8$ , the BVAR model with loose prior are suitable for negative collinearity level while the BVAR model with tight prior are suitable for positive collinearity levels.

Also for time series length and forecasting horizon of  $T = 16$ , the classical VAR model is preferred for all the collinearity levels except for  $\rho = -0.95$  and  $\rho = -0.9$  where the BVAR model with loose prior are preferred.

## 5. Conclusion and Recommendation

The results from this study revealed that the BVAR models were excellent for time series length of  $T = 8$  for all levels of collinearity while the classical VAR was effective for time series length of  $T = 16$  for all collinearity



**Table 4.** Ranks of the performances of the models for different time series length and forecasting horizons at different levels of collinearity.

COLLINEARITY LEVELS ( $\rho$ )	Models	T = 8		T = 16	
		RMSE	MAE	RMSE	MAE
-0.99	VAR(2)	N/A	N/A	1	1
	BVAR1	2	1	5	5
	BVAR2	1	2	4	4
	BVAR3	3	3	3	3
	BVAR4	4	4	2	2
-0.95	VAR(2)	N/A	N/A	3	3
	BVAR1	3	3	5	5
	BVAR2	4	4	4	4
	BVAR3	1	1	2	2
	BVAR4	2	2	1	1
-0.9	VAR(2)	N/A	N/A	3	1
	BVAR1	4	4	5	5
	BVAR2	3	3	4	4
	BVAR3	2	2	2	3
	BVAR4	1	1	1	2
-0.85	VAR(2)	N/A	N/A	1	1
	BVAR1	3	3	4	4
	BVAR2	4	4	5	5
	BVAR3	2	2	3	2
	BVAR4	1	1	2	3
-0.8	VAR(2)	N/A	N/A	1	1
	BVAR1	4	3	4	4
	BVAR2	3	4	5	5
	BVAR3	2	1	3	3
	BVAR4	1	2	2	2
0.8	VAR(2)	N/A	N/A	1	1
	BVAR1	1	1	2	2
	BVAR2	2	2	3	3
	BVAR3	3	3	5	5
	BVAR4	4	4	4	4
0.85	VAR(2)	N/A	N/A	1	1
	BVAR1	1	1	2	2
	BVAR2	2	2	3	3
	BVAR3	3	3	5	5
	BVAR4	4	4	4	4
0.9	VAR(2)	N/A	N/A	1	1
	BVAR1	2	1	2	2
	BVAR2	1	2	3	3
	BVAR3	3	3	5	5
	BVAR4	4	4	4	4

Continued

0.95	VAR(2)	N/A	N/A	1	1
	BVAR1	1	1	2	2
	BVAR2	2	2	3	3
	BVAR3	3	3	5	5
	BVAR4	4	4	4	4
0.99	VAR(2)	N/A	N/A	1	1
	BVAR1	1	1	2	2
	BVAR2	2	2	3	3
	BVAR3	3	3	4	4
	BVAR4	4	4	5	5

**Table 5.** Preferred models for different time series length and forecasting horizons at different levels of collinearity.

COLLINEARITY LEVELS ( $\rho$ )	T = 8		T = 16	
	RMSE	MAE	RMSE	MAE
-0.99	BVAR2	BVAR1	VAR(2)	VAR(2)
-0.95	BVAR3	BVAR3	BVAR4	BVAR4
-0.9	BVAR4	BVAR4	BVAR4	VAR(2)
-0.85	BVAR4	BVAR4	VAR(2)	VAR(2)
-0.8	BVAR4	BVAR3	VAR(2)	VAR(2)
0.8	BVAR1	BVAR1	VAR(2)	VAR(2)
0.85	BVAR1	BVAR1	VAR(2)	VAR(2)
0.9	BVAR2	BVAR1	VAR(2)	VAR(2)
0.95	BVAR1	BVAR1	VAR(2)	VAR(2)
0.99	BVAR1	BVAR1	VAR(2)	VAR(2)

levels except when  $\rho = -0.9$  and  $\rho = -0.95$ . Therefore, we recommended that for effective short term forecasting, the time series length, forecasting horizon and the collinearity level should be considered.

## Acknowledgements

We wish to thank TETFUND Abuja-Nigeria for sponsoring this research work. Our appreciation also goes to the Rector and the Directorate of Research, Conference and Publication of the Federal Polytechnic Bida for giving us this opportunity to undergo this research work.

## References

- [1] Sims, C.A. (1980) Macro Economics and Reality. *Econometrica*, **48**, 1-48. <http://dx.doi.org/10.2307/1912017>
- [2] Bischoff, C.W., Belay, H. and Kang, I-B. (2000) Bayesian VAR Forecasts Fail to Live up to Their Promise: For the Period 1981-1996. Judgementally Adjusted Large Structural Macroeconomic Models Forecast More Accurately than Bayesian VAR Models. *Business Economics*, **35**, 19-29.
- [3] Brandt, P.T. and Freeman, J.R. (2006) Advances in Bayesian Time Series Modeling and the Study of Politics: Theory, Testing, Forecasting and Policy Analysis. *Political Analysis*, **14**, 1-36. <http://dx.doi.org/10.1093/pan/mpi035>
- [4] Cottet, R and Smith, M. (2003) Bayesian Modelling and Forecasting of Intraday Electricity Load. *Journal of the American Statistical Association*, **98**, 839-849. <http://dx.doi.org/10.1198/016214503000000774>
- [5] Taylor, J.M. and Mcsharry, P.E. (2008) Short Term Load Forecasting Methods: An Evaluation Based on European Data. *IEEE Transactions on Power System*, **22**, 2213-2219. <http://dx.doi.org/10.1109/TPWRS.2007.907583>
- [6] Lacoviello, M. (2001) Short Term Forecasting: Projecting Italian GDP, One Quarter to Two Years Ahead. IMF Working Paper No. WP/01/109.
- [7] Karim, B., Szilard, B., Riccardo, C., Ard, D.R., Audrone, J., Pioyr, J., Antonio, R., Gerhard, R., Karsten, R. and Chris-

- tophe, V.N. (2008) Short Term Forecasting of GDP Using Large Monthly Data Sets. A Pseudo Real Time Forecast Evaluation Exercise. Occasional Paper Series No. 84, European Central Bank.
- [8] Caraiani, P. (2010) Forecasting Romanian GDP Using a BVAR Model. *Romanian Journal of Economic Forecasting*, **4**, 76-87
- [9] Godbout, C. and Lombardi, M.J. (2012) Short Term Forecasting of the Japanese Economy Using Factor Model. Working Paper Series No. 1428, European Central Bank.
- [10] Fahab, M.U. and Arbab, M.N. (2014) Factor Affecting Short Term Load Forecasting. *JCET*, **2**, 305-309.
- [11] Belousov, S.L. and Berkovich, L.V. (2015) Short Term Weather Forecasting. Environmental Structure and Function: Climate System—Volume I, EOLSS. <http://www.ecolss.net/Eolss-sampleAllchapter.aspx>
- [12] Gorr, W., Olligschlaeger, A. and Thompson, Y. (2003) Short Term Forecasting of Crime. *International Journal of Forecasting*, **19**, 579-594. [http://dx.doi.org/10.1016/S0169-2070\(03\)00092-X](http://dx.doi.org/10.1016/S0169-2070(03)00092-X)
- [13] Melihovs, A. and Rusakova, S. (2005) Short Term Forecasting of Economic Development in Latvia Using Business and Consumer Survey Data. Working Paper No. 4, RIGA.
- [14] Da, Z. and Warachka, M. (2011) The Disparity between Long Term and Short Term Forecasted Earnings Growth. *Journal of Financial Economics*, **100**, 424-442. <http://dx.doi.org/10.1016/j.jfineco.2010.10.015>
- [15] Johnson, T.L. (2009) Effect of Correlation and Identification Status on Methods of Estimating Parameters of System of Simultaneous Equations Model. Unpublished PhD Thesis, University of Ilorin, Ilorin.
- [16] Dormann, C.F., Elith, J., Bacher, S., Bachmann, C., Carl, G., Lagourcade, B., Leifao, P.J., Munkemiller, T., McClean, C., Osborne, P.E., Reineking, B., Schroder, B., Skidmore, A.K., Jurell, D. and Lautenbach, S. (2013) Collinearity: A Review of Methods to Deal with It and a Simulation Study Evaluating Their Performance. *Ecography*, **36**, 27-46. <http://dx.doi.org/10.1111/j.1600-0587.2012.07348.x>
- [17] Rama-Sastry, M.V. (1970) Some Limits in the Theory of Multicollinearity. *The American Statistician*, **24**, 39-40.
- [18] Gunasekara, F.I., Carter, K. and Blakely, T. (2008) Glossary for Econometrics and Epidemiology. *Journal of Epidemiology and Community Health*, **62**, 858-861. <http://dx.doi.org/10.1136/jech.2008.077461>
- [19] Mehta, J.S., Swamy, P.A.V.B. and Iyengar, N.S. (1993) Finite Sample Properties of a Measure of Multicollinearity. *Sankhya: The Indian Journal of Statistics B*, **55**, 353-371.
- [20] Buse, A. (1994) Brickmaking and the Collinear Arts: A Cautionary Tale. *The Canadian Journal of Economics*, **27**, 408-414. <http://dx.doi.org/10.2307/135754>
- [21] Clement, M.P. and Hendry, D.F. (1995) Macro-Economic Forecasting and Modelling. *The Economic Journal*, **105**, 1001-1013. <http://dx.doi.org/10.2307/2235166>
- [22] Hendry, D.F. (1997) The Econometrics of Macroeconomic Forecasting. *The Economic Journal*, **107**, 1330-1357. <http://dx.doi.org/10.1111/j.1468-0297.1997.tb00051.x>
- [23] Greenberg, E. and Parks, R.P. (1997) A Predictive Approach to Model Selection and Multicollinearity. *Journal of Applied Econometrics*, **12**, 67-75. [http://dx.doi.org/10.1002/\(SICI\)1099-1255\(199701\)12:1<67::AID-JAE427>3.0.CO;2-W](http://dx.doi.org/10.1002/(SICI)1099-1255(199701)12:1<67::AID-JAE427>3.0.CO;2-W)
- [24] Grewal, R., Cote, J.A. and Baumgartner, H. (2004) Multicollinearity and Measurement Error in Structural Equation Models: Implication for Theory Testing. *Marketing Science*, **23**, 519-529. <http://dx.doi.org/10.1287/mksc.1040.0070>
- [25] Friedman, L. and Wall, M. (2006) Graphical Views of Suppression and Multicollinearity in Multiple Linear Regression. *The American Statistician*, **59**, 127-136. <http://dx.doi.org/10.1198/000313005X41337>
- [26] Alabi, O.O., Ayinde, K. and Oyejola, B.A. (2007) Empirical Investigation of Effect of Multicollinearity on Type II Error Rates of the Ordinary Least Squares Estimators. *Journal of the Nigerian Statistical Association*, **19**, 50-57.
- [27] Ayyangar, L. (2007) Skewness, Multicollinearity, Heteroskedasticity—You Name It, Cost Data Have It! Solutions to Violations of Assumptions of Ordinary Least Regression Models Using SAS. SAS Global Forum, Orlando, 16-19 April 2007.
- [28] Ijomah, M.A. and Nduka, E.C. (2010) Selection in a Multicollinear Data Using Information Criteria. *Journal of the Nigerian Statistical Association*, **22**, 64-75.
- [29] Agunbiade, D.A. (2011) Effect of Multicollinearity and Sensitivity of the Methods in Simultaneous Equation Model. *Journal of Modern Mathematics and Statistics*, **5**, 9-12. <http://dx.doi.org/10.3923/jmmstat.2011.9.12>
- [30] Taylor, L.D. (2012) A Note on the Harmful Effects of Multicollinearity. [www.econ.arizona.edu/docs/Working\\_Papers/2012/Econ-WP-12-14.pdf](http://www.econ.arizona.edu/docs/Working_Papers/2012/Econ-WP-12-14.pdf)
- [31] Olanrewaju, S.O. (2013) Effect of Multicollinearity of Some Estimators of System of Regression Equations. *Proceedings of the 32nd Annual Conference of the Nigerian Mathematical Society*, Obafemi Awolowo University, Nigeria,

25-28 June 2013, 79.

- [32] Yahya, W.B. and Olaniran, O.R. (2013) Bayesian Linear Regression, Ridge Regression and Ordinary Least Squares Methods for modeling Collinear Data. *Proceedings of the 32nd Annual Conference of the Nigerian Mathematical Society*, Obafemi Awolowo University (OAU), 25-28 June 2013, 74.
- [33] Adenomom, M.O. and Oyejola, B.A. (2014) A Simulation Study of Effects of Collinearity on Forecasting of Bivariate Time Series Data. *Scholars Journal of Physics, Mathematics and Statistics*, **1**, 4-21.
- [34] Gujarati, D.N. (2003) Basic Econometrics. 4th Edition, The McGraw-Hill Co., New Delhi.
- [35] Lütkepohl, H. and Breitung, J. (1997) Impulse Response Analysis of Vector Autoregressive Processes. In: Heij, C., Schumacher, H., Hanzon, B. and Praagman, C., Eds., *System Dynamics in Economic and Financial Models*, John Wiley, New York.
- [36] Hamilton, J.D. (1994) Time Series Analysis. Princeton University Press, Princeton.
- [37] Zeller, A. (1962) An Efficient Method of Estimating Seemingly Unrelated Regression and Tests for Aggregation Bias. *Journal of the American Statistical Association*, **57**, 348-368. <http://dx.doi.org/10.1080/01621459.1962.10480664>
- [38] Sims, C.A. and Zha, T. (1998) Bayesian Methods for Dynamic Multivariate Models. *International Economic Review*, **39**, 949-968. <http://dx.doi.org/10.2307/2527347>
- [39] Brandt, P.T. and Freeman, J.R. (2009) Modeling Macro-Political Dynamics. *Political Analysis*, **17**, 113-142. <http://dx.doi.org/10.1093/pan/mpp001>
- [40] Sims, C.A. and Zha, T. (1999) Error Bands for Impulse Responses. *Econometrica*, **67**, 113-115. <http://dx.doi.org/10.1111/1468-0262.00071>
- [41] Diebold, F.X. and Mariano, R.S. (2002) Comparing Predictive Accuracy. *Journal of Business and Economic Statistics*, **20**, 134-144. <http://dx.doi.org/10.1198/073500102753410444>
- [42] Kadiyala, K.R. and Karlsson, S. (1997) Numerical Methods for Estimation and Inference in Bayesian VAR Models. *Journal of Applied Econometrics*, **12**, 99-132. [http://dx.doi.org/10.1002/\(SICI\)1099-1255\(199703\)12:2<99::AID-JAE429>3.0.CO;2-A](http://dx.doi.org/10.1002/(SICI)1099-1255(199703)12:2<99::AID-JAE429>3.0.CO;2-A)
- [43] Breheny, P. (2013) Wishart Priors. BST 701: Bayesian Modelling in Biostatistics. <http://web.as.uky.edu/statistics/users/pbreheny/701/S13/notes/3-28.pdf>
- [44] Brandt, P.T., Colaresi, M. and Freeman, J.R. (2008) Dynamic of Reciprocity, Accountability and Credibility. *Journal of Conflict Resolution*, **52**, 343-374. <http://dx.doi.org/10.1177/0022002708314221>
- [45] Merkin, J.H. and Needman, D.J. (1990) The Development of Travelling Waves in a Simple Isothermal Chemical System II. Cubic Autocatalysis with Quadratic and Linear Decay. *Proceedings: Mathematical & Physical Sciences*, **430**, 315-345. <http://dx.doi.org/10.1098/rspa.1990.0093>
- [46] Merkin, J.H. and Needman, D.J. (1991) The Development of Travelling Waves in a Simple Isothermal Chemical System IV. Quadratic Autocatalysis with Quadratic Decay. *Proceedings: Mathematical & Physical Sciences*, **434**, 531-554. <http://dx.doi.org/10.1098/rspa.1991.0112>
- [47] Worsley, K.J., Evans, A.C., Strother, S.C. and Tyler, J.L. (1991) A Linear Spatial Correlation Model with Applications to Positron Emission Tomography. *Journal of the American Statistical Association*, **86**, 55-67. <http://dx.doi.org/10.1080/01621459.1991.10475004>
- [48] Cooray, T.M.J.A. (2008) Applied Time series Analysis and Forecasting. Narosa Publishing House, New Delhi.
- [49] Gilbert, P. (2009) Brief User's Guide: Dynamic Systems Estimation (DSE). [www.bank-banque-canada.ca/pgilbert](http://www.bank-banque-canada.ca/pgilbert)
- [50] Pfaff, B. (2008) VAR, SVAR and SVEC Models: Implementation within R Package Vars. *Journal of Statistical Software*, **27**, 1-32. <http://dx.doi.org/10.18637/jss.v027.i04>
- [51] Brandt, P.T. (2012) Markov-Switching, Bayesian Vector Autoregression Models-Package "MSBVAR". The R Foundation for Statistical Computing.