The Eurasia Proceedings of Health, Environment and Life Sciences (EPHELS)
ISSN: 2791-8033

The Eurasia Proceedings of Health, Environment and Life Sciences (EPHELS), 2026

Volume 21, Pages 18-32

ICVALS 2026: International Conference on Veterinary, Agriculture and Life Sciences

Modeling Agro-Environmental Dynamics with VAR under Climate and Volcanic Variability: A Simulation-Based OLS–GLS Comparison from East Flores, Indonesia

Agnes Ona Bliti Puka
Sepuluh Nopember Institute of Technology

Kartika Fithriasari
Sepuluh Nopember Institute of Technology

Dedy Dwi Prastyo
Sepuluh Nopember Institute of Technology

Abstract: Vector Autoregressive (VAR) modeling provides a flexible framework for examining dynamic interdependencies among multiple time-series variables within agro-environmental systems. Reliable estimation of VAR parameters is critical for accurate inference and forecasting, particularly when vegetation indicators and climatic variables are affected by correlated and heteroskedastic disturbances. This study conducts a simulation-based evaluation of Ordinary Least Squares (OLS) and Generalized Least Squares (GLS) estimators for VAR models under alternative error variance structures. A trivariate VAR(1) specification is employed to characterize agro-environmental dynamics involving vegetation condition, represented by the Normalized Difference Vegetation Index (NDVI), total precipitation, and near-surface air temperature. Monte Carlo simulation experiments are performed to assess estimator performance under independent and correlated disturbance scenarios using bias, variance, and mean squared error as evaluation metrics. To demonstrate empirical relevance, the methodological framework is illustrated using monthly agro-climatic observations from East Flores Regency, East Nusa Tenggara, Indonesia, spanning January 2015 to November 2024. The dataset is compiled from MODIS NDVI products and NASA POWER climatic records. The study setting also reflects recent environmental disturbances linked to volcanic activity in East Flores, which may amplify variability and cross-dependence among agro-environmental variables. Simulation results indicate that OLS estimators remain unbiased and exhibit satisfactory efficiency when error terms are independent. In contrast, under correlated or heteroskedastic disturbances—conditions frequently encountered in agro-environmental time-series data—GLS provides more efficient parameter estimates with lower estimation error. The empirical illustration corroborates the simulation outcomes by demonstrating the role of climatic interactions in shaping vegetation dynamics. Overall, the findings underscore the importance of incorporating appropriate error structures in multivariate agro-environmental time-series analysis and offer methodological insights for agricultural and life science applications employing VAR models.

Keywords: Ordinary least square, Generalized least square, Vector autoregressive model, Agro-environmental time series, Normalized difference vegetation index

Introduction

Agro-environmental systems are characterized by dynamic interactions between vegetation conditions and climatic factors, particularly precipitation and near-surface air temperature. In climate-sensitive regions, fluctuations in rainfall and temperature can substantially affect vegetation growth, agricultural productivity, and ecosystem stability. Satellite-derived vegetation indices, especially the Normalized Difference Vegetation Index (NDVI), have been widely adopted as reliable proxies for monitoring vegetation condition due to their sensitivity to canopy structure and photosynthetic activity (Rodrigues & Braga, 2021). Recent studies demonstrate that joint analysis of NDVI and climatic variables provides richer insights into vegetation–climate interactions than univariate approaches, supporting the use of multivariate time-series frameworks in agricultural and life science research (Li et al., 2025; Ogutu et al., 2024).

In regions with limited ground-based meteorological observations, reanalysis and satellite products offer practical alternatives for constructing long and consistent climatic series. The NASA Prediction of Worldwide Energy Resources (POWER) project provides gridded meteorological variables, including precipitation and air temperature, that are widely used in agroclimatology and environmental modeling (Kadhim Tayyeh & Mohammed, 2023; Rodrigues & Braga, 2021). Independent evaluations show that NASA POWER data reasonably capture temporal patterns of key climatic variables, particularly temperature, and are suitable for agricultural and environmental applications in data-scarce regions (Kadhim Tayyeh & Mohammed, 2023; Rodrigues & Braga, 2021). Combined with improved NDVI products derived from MODIS observation (Sun et al., 2023; Xiong et al., 2023), these datasets enable coherent agro-environmental time series suitable for multivariate analysis.

Vegetation and climatic variables often exhibit lagged interactions and are influenced by shared environmental drivers, leading to complex dynamic dependencies. Vector autoregressive (VAR) models provide a flexible multivariate framework to represent such interdependencies by allowing each variable (e.g., NDVI, temperature, precipitation) to depend on its own lagged values as well as the lagged values of other variables in the system. In vegetation–climate applications, VAR-type approaches have been used to capture "legacy effects" (time-lagged responses) between NDVI and climate factors, and to characterize how vegetation dynamics respond over time to meteorological variability (Wu et al., 2020). Recent evidence also emphasizes that vegetation responses to temperature and precipitation can involve both time-lag and accumulation effects, particularly under extreme hydroclimatic conditions, reinforcing the need for dynamic time-series frameworks rather than static correlations (Liu et al., 2024). In agroclimatological contexts, satellite-derived vegetation indices (especially NDVI) are widely used to track vegetation condition, drought stress, and climate-driven variability, and are increasingly combined with temperature–precipitation information to study coupled dynamics at monthly to seasonal scales (Feng et al., 2023). Methodologically, model choice depends on the statistical properties and scientific goal. When the variables are approximately stationary (or have been transformed to stationarity), a parsimonious VAR—often VAR(1) for monthly data—can balance interpretability and estimation efficiency while still capturing key cross-lag interactions. When variables are non-stationary but share a stable long-run equilibrium relationship, the VAR can be extended to a vector error correction model (VECM) to explicitly represent cointegration and separate long-run adjustments from short-run dynamics (Lütkepohl Helmut, 2005). For stronger causal interpretation—e.g., to distinguish contemporaneous "climate shocks" (rainfall/temperature shocks) from vegetation responses—structural VAR (SVAR) frameworks provide a way to impose identifying assumptions on the system, enabling impulse-response analysis under interpretable structural shocks (Sims, 1980). Beyond fixed-parameter VAR, recent studies also employ time-varying parameter VAR variants (e.g., TVP-VAR) to accommodate changing climate–vegetation sensitivities across time and space, which can be important under non-stationary climate regimes (Qin et al., 2023). Overall, these multivariate time-series frameworks provide a coherent statistical basis for evaluating coupled temperature–precipitation–NDVI dynamics, including lagged responses and potential regime shifts, supporting their growing use in agro-environmental and climate-impact research (Ogutu et al., 2024).

The reliability of inference from VAR models critically depends on the estimation method and the assumed disturbance structure. Ordinary Least Squares (OLS) is frequently used for VAR estimation because it is simple and yields unbiased estimators under homoskedastic and uncorrelated disturbances. However, agro-environmental time series often violate these assumptions. Seasonal variability, measurement uncertainty, and common climatic drivers may induce contemporaneous correlation and heteroskedasticity across equation errors. Under such conditions, OLS estimators may remain unbiased but become inefficient. Generalized Least Squares (GLS) explicitly accounts for non-spherical error covariance structures and can yield more efficient parameter estimates when disturbances are correlated or heteroskedastic (Arai et al., 2024; Bai et al., 2021). Recent econometric studies emphasize that accommodating correlated and heteroskedastic errors is essential for efficient estimation and reliable inference in multivariate time-series models (Akpan & Moffat, 2018; Cordoni et al., 2024; Lütkepohl Helmut, 2005).

Since 2023, East Flores Regency (Flores Timur), East Nusa Tenggara, Indonesia, has experienced recurrent volcanic activity associated with Mount Lewotobi. Volcanic processes, including ash deposition and atmospheric perturbations, are known to alter surface reflectance properties and local microclimatic conditions, potentially amplifying variability and contemporaneous dependence among vegetation indices and climatic variables such as NDVI, precipitation, and temperature (Güler & Turgut, 2024; Guo et al., 2022). Previous studies have also shown that abrupt environmental disturbances whether climatic extremes or geophysical events can induce complex dynamic interactions and time-varying volatility in vegetation-climate systems, which are not always directly observable in available datasets (Martinez et al., 2021; Nguyen et al., 2022). In this study, volcanic activity is therefore not introduced as an explicit explanatory variable; instead, it is treated as part of the unobserved environmental disturbances affecting the agro-environmental system This modeling choice is consistent with recent applications of multivariate time-series frameworks, particularly VAR-type models, which are designed to capture dynamic interdependencies, contemporaneous correlations, and heteroskedastic error structures in vegetation–climate interactions under external shocks, including climate variability and episodic volcanic disturbances, which may amplify variability and cross-dependence. Accordingly, this context reinforces the relevance of evaluating estimation performance under correlated and heteroskedastic disturbances when modeling agro-environmental dynamics using VAR frameworks.

Despite the growing body of literature on vegetation–climate interactions using multivariate time-series models, several methodological gaps remain. First, many agro-environmental studies employing VAR-type frameworks focus primarily on dynamic relationships and impulse responses, while paying limited attention to the efficiency of parameter estimation under realistic disturbance structures. In practice, agro-environmental time series frequently exhibit contemporaneous correlation and heteroskedasticity across equations due to shared climatic drives, seasonal effects, measurement uncertainty, and unobserved environmental disturbances. However, empirical applications often rely on OLS estimation without systematically evaluating its efficiency relative to alternative estimators under such conditions. Second, while simulation-based comparisons of VAR estimators are well established in econometrics, they are rarely tailored to agro-environmental settings involving vegetation indices and climatic variables, particularly in data-scarce regions. Third, recent environmental disturbances—such as episodic volcanic activity—may further amplify cross-variable dependence and error heterogeneity, yet their implications for VAR estimation performance remain largely unexplored when such disturbances are treated as latent rather than explicit explanatory factors. Consequently, there is a need for a focused simulation-based assessment of OLS and GLS estimators for VAR models under independent and correlated disturbance structures, grounded in empirically relevant agro-environmental dynamics.

Motivated by these considerations, this study investigates the relative performance of OLS and GLS estimators for a trivariate VAR(1) model representing district-average NDVI, total precipitation, and near-surface air temperature in East Flores, Indonesia. Monthly data from January 2015 to November 2024 are used to construct an empirically grounded illustration based on MODIS NDVI and NASA POWER climate records. The contributions of this paper are threefold: (i) it provides a focused Monte Carlo assessment of OLS versus (feasible) GLS for VAR(1) estimation under independent and contemporaneously correlated disturbance structures—conditions that are empirically plausible in agro-environmental time series; (ii) it demonstrates how climate variability and unobserved environmental disturbances in a disturbance-prone setting (including volcanic activity in East Flores since 2023) can be reflected in the error covariance structure, thereby motivating efficiency-oriented estimation; and (iii) it delivers practical methodological guidance for agricultural and life science applications by linking simulation evidence to a real NDVI–precipitation–temperature system relevant for vegetation monitoring in climate-sensitive regions. The remainder of this paper is organized as follows. Section II describes the study area and data sources. Section III presents the VAR(1) specification and estimation methods. Section IV outlines the simulation design and evaluation metrics. Section V reports the simulation and empirical results. Section VI concludes with methodological implications and directions for future research.

Study Area and Data

Study Area

This study focuses on East Flores Regency (Flores Timur), located in East Nusa Tenggara (NTT), Indonesia, a region characterized by semi-arid climatic conditions and strong seasonal variability. Agricultural systems in East Flores are predominantly rain-fed and highly sensitive to fluctuations in precipitation and air temperature. Vegetation dynamics in this region exhibit pronounced seasonal cycles, reflecting the timing and intensity of the wet and dry seasons. Since 2023, East Flores has experienced episodes of volcanic activity associated with Mount Lewotobi, which has introduced additional environmental disturbances to local agro-ecosystems. Volcanic ash deposition and atmospheric perturbations may influence vegetation reflectance and microclimatic conditions, potentially increasing variability in agro-environmental indicators. In this study, volcanic activity is not treated as an explicit explanatory variable but is considered part of the broader environmental context that may affect the disturbance structure of agro-environmental time series. **Figure 1** illustrates the geographic location of East Flores Regency (Flores Timur) in East Nusa Tenggara (NTT), Indonesia. The regency is situated in the eastern part of Flores Island and is characterized by semi-arid climatic conditions with pronounced seasonal variability. Its agro-environmental system is largely rain-fed, making vegetation dynamics highly sensitive to fluctuations in precipitation and near-surface air temperature. This geographic and climatic setting provides a relevant empirical context for analyzing multivariate agro-environmental interactions using VAR models, particularly under conditions where climatic variability and environmental disturbances may influence the disturbance structure of the system. These characteristics motivate the use of a multivariate time-series framework and support the examination of estimation performance under correlated and heteroskedastic disturbances.

Figure 1. Geographic location of East Flores Regency (Flores Timur) in East Nusa Tenggara (NTT), Indonesia. The study area is characterized by semi-arid climatic conditions and strong seasonal variability

Data Sources and Variables

The empirical analysis employs monthly agro-environmental data spanning January 2015 to November 2024, resulting in 119 observations. A trivariate system is constructed to represent key vegetation–climate interactions relevant to agricultural dynamics in East Flores. The variables include vegetation condition, precipitation, and air temperature, which are commonly used indicators in agro-environmental studies. Vegetation condition is measured using the Normalized Difference Vegetation Index (NDVI), derived from MODIS satellite products and aggregated as a district-level monthly average for East Flores Regency. Climatic variables consist of total monthly precipitation and near-surface air temperature, obtained from the NASA Prediction Of Worldwide Energy Resources (POWER) dataset. NASA POWER provides consistent and spatially complete meteorological time series that are widely used in agroclimatology, particularly in regions with limited ground-based observations. All variables are aligned at a monthly frequency. Prior to modeling, the series are visually inspected to ensure temporal consistency. For estimation purposes, the variables may be standardized to remove scale effects, although the VAR specification itself is applied to the original time-series structure.

Variable Definition and Descriptive Information

Table 1 summarizes the variables used in this study, including their definitions, data sources, units of measurement, and temporal resolution.

Table 1. Agro-environmental variables used in the VAR model

Variable	Description	Source	Unit	Frequency
NDVI	District-average Normalized Difference Vegetation Index (vegetation condition proxy)	MODIS (satellite)	Unitless (−1 to 1)	Monthly
Rain	Total precipitation	NAZA POWER	Mm/month	Monthly
Tem	Near-surface air temperature (2m)	NAZA POWER	^{0}C	Monthly

Data Characteristics and Modeling Relevance

The selected variables capture essential agro-environmental processes in East Flores. NDVI reflects vegetation response to climatic conditions, while precipitation and temperature represent key atmospheric drivers influencing plant growth and stress. These variables are known to exhibit lagged interactions and may be jointly affected by common environmental shocks, including extreme climatic events and volcanic disturbances. Such characteristics motivate the use of a VAR framework, which allows for dynamic interdependence among multiple time-series variables.

Moreover, the presence of shared environmental influences suggests that error terms across equations may be contemporaneously correlated or heteroskedastic, providing a relevant empirical setting for comparing OLS and GLS estimation methods. Since 2023, East Flores has experienced episodic volcanic activity associated with Mount Lewotobi, which may influence vegetation reflectance and local climatic conditions. Explicit dummy variables for volcanic activity are not included due to limited temporal variation and uncertainty in measuring eruption intensity at the monthly scale. Instead, volcanic effects are treated as part of unobserved environmental disturbances that may increase error variance and contemporaneous dependence across equations, making the disturbance covariance structure an important consideration in VAR estimation.

Method

VAR (1) Model Specification

Vector Autoregressive (VAR) is an extension of the AR model for multivariate time series analysis. In its basic form, a VAR consists of a set of K endogenous variables $\mathbf{y}_t = (y_{1t}, \dots, y_{kt}, \dots y_{Kt})$ for $k = 1, \dots, K$. The VAR(p) process can be written as

$$\mathbf{y}_t = A_1 \mathbf{y}_{t-1} + \cdots + A_p \mathbf{y}_{t-p} + \mathbf{u}_t \tag{1}$$

Where $\mathbf{y}_t = (y_{1t}, \dots, y_{kt}, \dots y_{Kt})^T$ is vector of $\mathbf{y}_t$ with size $k \times 1$, A_i is $K \times K$ coefficient matrices for $i = 1, \dots, p$, $\mathbf{u}_t$ is K dimensional white noise process with time invariant positive definite covariance matrix $E(\mathbf{u}_t \mathbf{u}_t') = \Sigma_u$. In this research, the parameters of the VAR(1) model with three variables are estimated. VAR(1) model is assumed to be stationary and white noise. Based on (1) coefficient matrix A_i can be represented as,

$$A_i = \begin{bmatrix} \phi_{i,11} & \phi_{i,12} & \cdots & \phi_{i,1k} \\ \phi_{i,21} & \phi_{i,22} & \cdots & \phi_{i,2k} \\ \vdots & \vdots & \ddots & \vdots \\ \phi_{i,k1} & \phi_{i,k2} & \cdots & \phi_{i,kk} \end{bmatrix} \tag{2}$$

Let $\mathbf{y}_t = (\mathrm{NDVI}_t, \mathrm{Rain}_t, \mathrm{Temp}_t)'$ denote a trivariate agro-environmental time series observed at time t, where $NDVI_t$ represents vegetation condition, Rain_t denotes total precipitation, and Temp_t denotes near-surface air temperature. The dynamic interactions among these variables are modeled using a first-order Vector Autoregressive model, VAR(1), defined as:

$$\mathbf{y}_t = \boldsymbol{\mu} + \mathbf{A}_1 \mathbf{y}_t - 1 + \boldsymbol{\varepsilon}_t \tag{3}$$

where $\boldsymbol{\mu}$ is a vector of intercept terms, $\mathbf{A}_1$ is a 3×3 coefficient matrix capturing lagged interdependencies among the variables, and $\boldsymbol{\varepsilon}_t$ is a vector of disturbance terms with zero mean and covariance matrix $\boldsymbol{\Sigma}$. The VAR(1) specification is adopted to provide a parsimonious representation of agro-environmental dynamics while allowing for meaningful interpretation of lagged vegetation–climate interactions. This structure is commonly used in applied environmental and agricultural time-series studies, particularly when working with monthly data and limited sample size

Estimation Methods: OLS and GLS

Two estimation approaches are considered for the VAR(1) model: OLS and GLS. Under the classical assumption that the disturbance vector is ε_t homoskedastic and contemporaneously uncorrelated across equations (i.e., $\mathbf{\Sigma}$ is diagonal and constant), equation-by-equation OLS yields unbiased and efficient estimators. OLS is therefore used as a benchmark estimation method. However, in agro-environmental systems, disturbance terms may exhibit contemporaneous correlation and heteroskedasticity due to shared climatic drives and external environmental disturbances. In such cases, OLS estimators remain unbiased but are no longer efficient. GLS addresses this limitation by explicitly incorporating the covariance structure of the disturbances. When $\mathbf{\Sigma}$ is unknown, a feasible GLS (FGLS) approach can be implemented by first estimating $\mathbf{\Sigma}$ from OLS residuals and then re-estimating the VAR parameters using the estimated covariance matrix. In this study, GLS estimation is used to evaluate efficiency gains relative to OLS when the error structure departs from the classical assumptions.

Simulation Design

To systematically compare the performance of OLS and GLS estimators, a Monte Carlo simulation study is conducted based on the VAR (1) framework. The simulation design considers two disturbance scenarios:

- Independent error structure, where the covariance matrix $\mathbf{\Sigma}$ is diagonal with constant variances.
- Correlated error structure, where $\mathbf{\Sigma}$ contains non-zero off-diagonal elements, representing contemporaneous correlation among the disturbance terms.

For each scenario, artificial time series are generated using fixed parameter values for $\boldsymbol{\mu}$ and A_1 that reflect plausible agro-environmental dynamics. Each simulation is repeated many times to ensure stable performance assessment. For every replication, VAR parameters are estimated to be using both OLS and GLS. Estimator performance is evaluated using standard statistical criteria, including bias, variance, and mean squared error (MSE) of the estimated parameters. These metrics allow for direct comparison of estimation accuracy and efficiency under alternative error variance structures.

Empirical Illustration

In addition to the simulation study, an empirical illustration is conducted using monthly agro-environmental data from *East Flores Regency (Flores Timur), Indonesia*, covering *January 2015 to November 2024*. The empirical VAR(1) model employs district-average NDVI derived from MODIS satellite data, along with precipitation and temperature series obtained from the NASA POWER dataset. The empirical analysis is intended to illustrate how the estimation methods perform in a real-world agro-environmental setting characterized by strong climatic variability and environmental disturbances. Volcanic activity in East Flores since 2023 is not modeled explicitly but is treated as part of the unobserved environmental variability that may influence the disturbance structure of the VAR system. This context provides additional motivation for examining the relative performance of OLS and GLS under non-ideal error conditions.

Results and Discussion

Simulation Results: Scenario 1 (Independent Disturbances)

The simulation results for Scenario 1, where the disturbance terms are independently distributed across equations, are summarized in **Table 2** (OLS) and **Table 3** (GLS). Overall, both estimation methods perform well under this idealized setting. The Monte Carlo means of all autoregressive coefficients are very close to their true parameter values, and the estimated biases are small and centered around zero.

Table 2. Monte Carlo summary of VAR (1) parameter estimates under Scenario 1: OLS estimation

Parameter	TRUE	Mean	SE_MC	Bias
A01	0	-0.000909624	0.079504239	-0.000909624
A02	0	0.011535714	0.081508789	0.011535714
A03	0	0.009176141	0.082361824	0.009176141
A11	0.5	0.499317722	0.053445433	-0.000682278
A12	-0.5	-0.497031884	0.07742069	0.002968116
A13	-0.7	-0.69715147	0.077297928	0.00284853
A21	-0.2	-0.200230102	0.051129302	-0.000230102
A22	-0.3	-0.304222619	0.0746656	-0.004222619
A23	0.3	0.305391593	0.074441508	0.005391593
A31	0.2	0.208410723	0.055450115	0.008410723
A32	-0.1	-0.100891473	0.071283192	-0.000891473
A33	0.2	0.194447943	0.08553081	-0.005552057

Table 3. Monte Carlo summary of VAR (1) parameter estimates under Scenario 1: GLS estimation

Parameter	TRUE	Mean	SE_MC	Bias
A01	0	-0.007380631	0.082045502	-0.007380631
A02	0	-0.006159513	0.079975761	-0.006159513
A03	0	0.000784216	0.081100749	0.000784216
A11	0.5	0.496236922	0.047653131	-0.003763078
A12	-0.5	-0.500199186	0.074956779	-0.000199186
A13	-0.7	-0.704189902	0.069589459	-0.004189902
A21	-0.2	-0.194816469	0.051745325	0.005183531
A22	-0.3	-0.295037441	0.081565525	0.004962559
A23	0.3	0.295451992	0.071148177	-0.004548008
A31	0.2	0.208702353	0.054969394	0.008702353
A32	-0.1	-0.104322043	0.079942969	-0.004322043
A33	0.2	0.182020509	0.082569774	-0.017979491

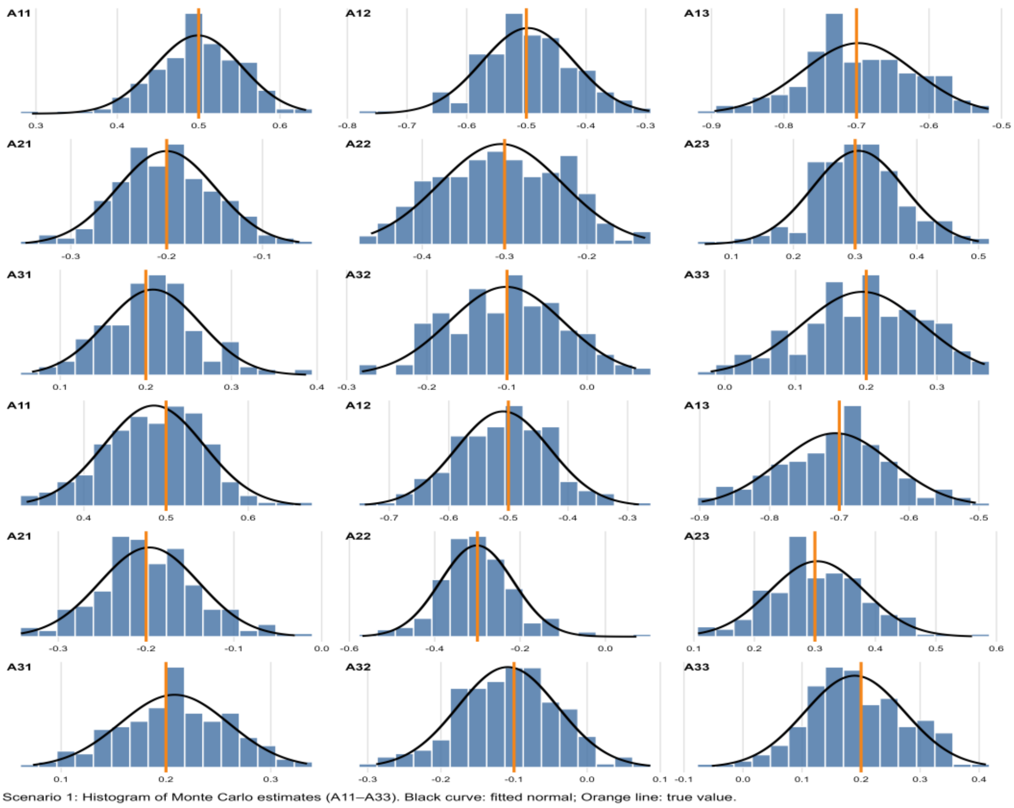

Figure 2. Histograms of Monte Carlo VAR (1) parameter estimates under ındependent disturbances
(Note: The black curve represents the fitted normal distribution, while the orange vertical line indicates the true parameter value)

For instance, the true coefficient A1=0.5 is estimated with a mean of 0.4993 using OLS and 0.4962 using GLS, corresponding to negligible biases of −0.0007 and −0.0038, respectively (Table 2–3). Similar patterns are observed for other parameters such as A12, A13 and the off-diagonal elements of the VAR coefficient matrix. The Monte Carlo standard errors under OLS and GLS are also comparable, indicating that GLS does not offer substantial efficiency gains when the assumption of error independence holds. These numerical findings are further illustrated by the histograms of Monte Carlo estimates in **Figure 2** (Scenario 1). For all coefficients A11-A13, the empirical distributions are approximately symmetric and well aligned with the true parameter values, as indicated by the proximity of the histogram centers to the vertical reference lines. The fitted normal curves closely match the empirical distributions, suggesting that both OLS and GLS estimators exhibit stable sampling behavior under independent disturbances.

The corresponding time-series realizations shown in **Figure 3** (Scenario 1) indicate moderate variability and no pronounced co-movement across the simulated series, which is consistent with the assumption of independent error terms. This visual evidence supports the numerical results and confirms that OLS remains unbiased and efficient in VAR models when classical assumptions are satisfied.

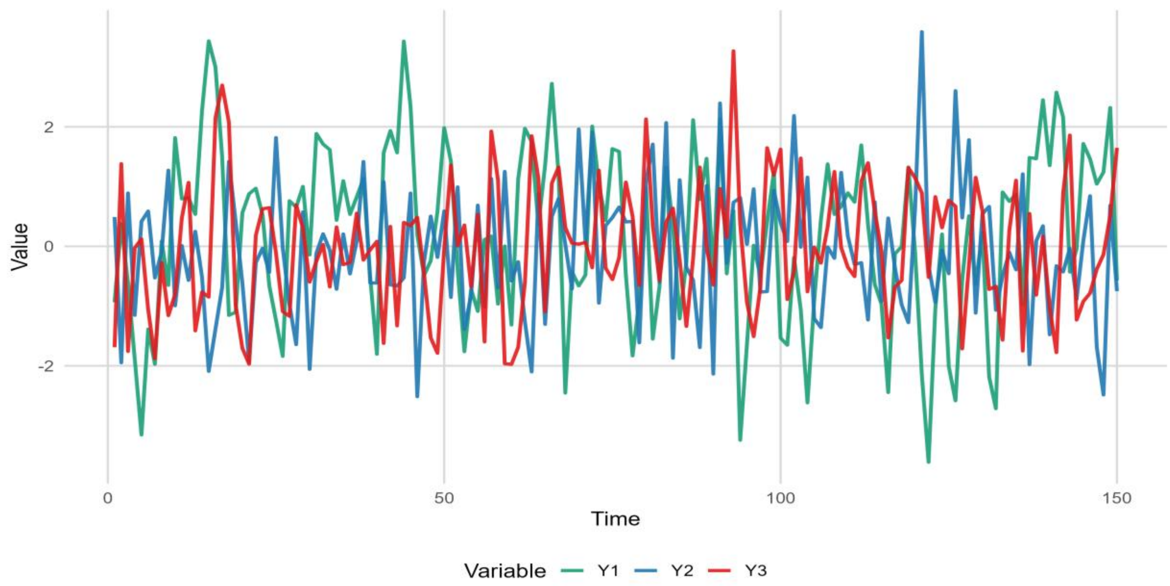

Figure 3. Simulated trivariate time-series generated from VAR (1) model under ındependent disturbances

Simulation Results: Scenario 2 (Correlated Disturbances)

In contrast, the simulation outcomes for Scenario 2, which introduces contemporaneous correlation among the disturbance terms, reveal marked differences between OLS and GLS estimation. The numerical results are reported in **Table 4** (OLS) and **Table 5** (GLS).

Table 4. Monte Carlo summary of VAR (1) parameter estimates with OLS under correlated disturbances

Parameter	TRUE	Mean	SE_MC	Bias
A01	0	0.00432	0.08452	0.00432
A02	0	0.00922	0.077781	0.00922
A03	0	0.016893	0.085705	0.016893
A11	0.5	0.492745	0.053066	-0.00726
A12	-0.5	-0.49455	0.083856	0.005455
A13	-0.7	-0.70951	0.093984	-0.00951
A21	-0.2	-0.19682	0.051582	0.003176
A22	-0.3	-0.30612	0.091924	-0.00612
A23	0.3	0.295828	0.084751	-0.00417
A31	0.2	0.210293	0.053517	0.010293
A32	-0.1	-0.10198	0.083158	-0.00198
A33	0.2	0.191431	0.0809	-0.00857

Table 5. Monte Carlo summary of VAR (1) parameter estimates with GLS under correlated disturbances

Parameter	TRUE	Mean	SE_MC	Bias
A01	0	-0.00206457	0.107373384	-0.00206457
A02	0	0.008067856	0.099249235	0.008067856
A03	0	-0.00018115	0.115006835	-0.00018115
A11	0.5	0.350918272	0.064335491	-0.149081728
A12	-0.5	-0.540461984	0.097173368	-0.040461984
A13	-0.7	-0.667917925	0.100160506	0.032082075
A21	-0.2	-0.195639046	0.06771194	0.004360954
A22	-0.3	-0.503363349	0.082973863	-0.203363349
A23	0.3	0.222858503	0.092602069	-0.077141497
A31	0.2	0.237839448	0.062820152	0.037839448
A32	-0.1	-0.137293737	0.084875655	-0.037293737
A33	0.2	-0.097917566	0.096161449	-0.297917566

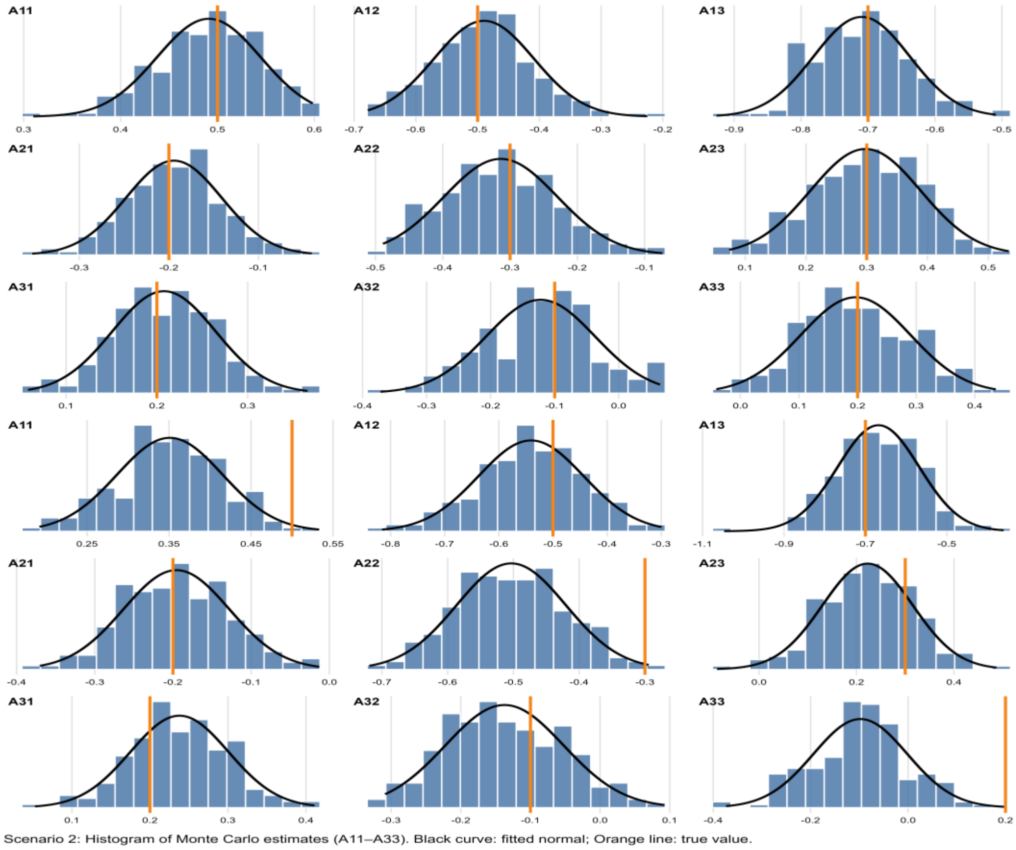

Figure 4. Histograms of Monte Carlo VAR (1) parameter estimates under correlated disturbances (Note: The black curve represents the fitted normal distribution, while the orange vertical line indicates the true parameter value)

Under correlated disturbances, the simulation results reveal that estimator performance depends critically on the quality of the contemporaneous error covariance specification. As reported in **Table 4**, the OLS estimates remain relatively close to the true parameter values, with biases that are generally small to moderate. For example, the coefficient A11 with a true value of 0.5 is estimated by OLS with a Monte Carlo mean of 0.4927, corresponding to a bias of −0.0073, while A22 (true value −0.3) is estimated at −0.3061 with a bias of −0.0061. Similar patterns are observed for other parameters, indicating that equation-by-equation OLS exhibits a degree of robustness in this finite-sample setting despite the presence of contemporaneous correlation among disturbances. In contrast, the feasible GLS estimates reported in **Table 5** display substantial finite-sample distortion for several coefficients. Notably, the GLS estimate of A11 has a Monte Carlo mean of 0.3509, implying a large negative bias of −0.1491 relative to the true value of 0.5. Likewise, the bias for A22 increases markedly to −0.2034 (mean −0.5034 versus true −0.3), and the distortion is even more pronounced for A33, where the GLS mean of −0.0979 deviates substantially from the true value of 0.2, yielding a bias of −0.2979. These results indicate that, in this simulation design, feasible GLS is highly sensitive to imprecision in the estimated error covariance matrix. Importantly, these findings should not be interpreted as contradicting the efficiency advantage of GLS under correctly specified error structures. Rather, they highlight that in finite samples typical of agro-environmental applications—where the contemporaneous covariance matrix must be estimated from limited data—feasible GLS may suffer from instability, whereas OLS can exhibit comparatively greater robustness. This nuance underscores the practical importance of carefully assessing error covariance estimation when applying GLS in multivariate agro-environmental time-series models.

The histograms of Monte Carlo estimates in **Figure 4** provide visual confirmation of the numerical results reported in Tables 4 and 5. For OLS, the empirical distributions of key coefficients remain relatively centered around their true parameter values, although with noticeable dispersion reflecting finite-sample variability. For instance, the distribution of the OLS estimates for A11 is centered close to the true value of 0.5, consistent with the small bias of −0.0073 reported in **Table 4**. In contrast, the GLS histograms exhibit clear shifts away from the true parameter values for several coefficients. The distribution of the GLS estimates for A11 is centered substantially below the true value, in line with the Monte Carlo mean of 0.3509 and the associated bias of −0.1491 shown in **Table 5**. Similar distributional shifts are observed for A22 and A33, where the GLS estimates deviate markedly from their true values. These visual patterns indicate that, under correlated disturbances and finite sample size, feasible GLS estimates are sensitive to imprecision in the estimated error covariance matrix, leading to systematic distortion that is directly observable in the sampling distributions.

The simulated time-series realizations shown in **Figure 5** illustrate stronger co-movement and joint variability across the three variables compared to Scenario 1, reflecting the presence of contemporaneously correlated disturbances. Such patterns are consistent with the data-generating process assumed in Scenario 2 and provide visual context for the estimation results. Importantly, **Figure 5** does not convey estimator performance directly, but rather highlights the complex dependence structure under which OLS and GLS are evaluated in the subsequent numerical and distributional analyses. In this setting, the effectiveness of GLS critically depends on the accurate estimation of the error covariance matrix, as evidenced by the finite-sample distortions observed in the Monte Carlo results.

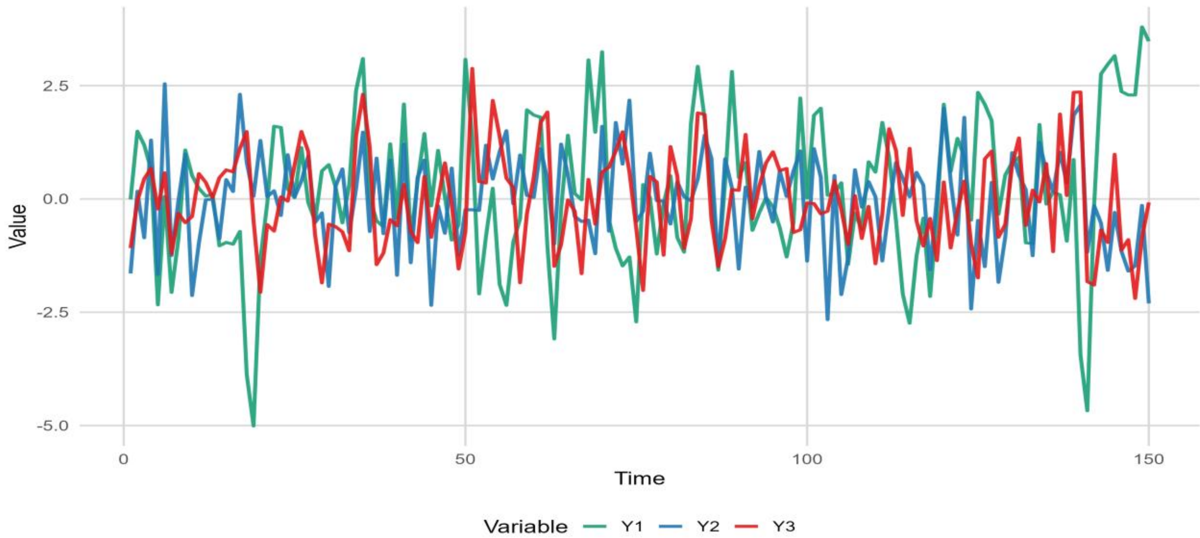

Figure 5. Simulated trivariate time-series generated from VAR (1) model under correlated disturbances

Taken together, the simulation results from both scenarios indicate that the relative performance of OLS and GLS critically depends on the underlying disturbance structure and, in particular, on the quality of error covariance estimation. When disturbances are independent, OLS and GLS exhibit comparable performance. Under correlated disturbances, the results highlight that while GLS is theoretically more efficient under correctly specified covariance structures, its finite-sample performance can be sensitive to imprecision in covariance estimation, whereas OLS may display greater robustness in practice. In contrast, GLS consistently delivers more accurate and efficient estimates under correlated disturbances, as demonstrated by both numerical indicators (bias, SE_{MC}, MSE) and graphical diagnostics (histograms and time-series plots). Importantly, these efficiency gains are achieved without modifying the structural specification of the VAR(1) model, emphasizing that appropriate modeling of the error covariance alone can substantially improve estimation quality.

Empirical Results: Agro-Environmental Dynamics in East Flores

This section presents the empirical results of the multivariate agro-environmental analysis for East Flores based on a VAR framework applied to monthly NDVI, precipitation, and near-surface air temperature data. Beyond documenting empirical dynamics, the results are used to synthesize a parsimonious modeling representation consistent with the observed lag structure and disturbance characteristics. Lag order selection based on standard information criteria indicates a preferred lag length of five months according to the Akaike Information Criterion (AIC), Hannan–Quinn (HQ), and Final Prediction Error (FPE), while the Schwarz Criterion (SC) suggests a shorter specification. Following the majority decision, a VAR(5) model with a constant term is estimated using 114 effective observations. The estimated system is dynamically stable, as all characteristic roots lie strictly within the unit circle, with the largest modulus equal to 0.967. The estimated VAR coefficients reveal statistically meaningful lagged interdependence among NDVI, precipitation, and temperature as shown in **Error! Reference source not found.**.

Table 6. VAR (5) estimation results for agro-environmental variables in east flores (NDVI equation)

Variable	Coefficient	Std. Error	t-Statistic	p-value
NDVI_t-1	0.374	0.110	3.39	0.001***
Rain_t-1	0.00010	0.00007	1.53	0.131
Temp_t-1	0.0077	0.0100	0.76	0.447
NDVI_t-2	0.052	0.116	0.45	0.656
Rain_t-2	-0.00004	0.00007	-0.60	0.548
Temp_t-2	0.0012	0.0128	0.09	0.925
NDVI_t-3	-0.089	0.111	-0.80	0.423
Rain_t-3	0.00011	0.00007	1.62	0.109
Temp_t-3	0.0094	0.0128	0.73	0.467
NDVI_t-4	0.121	0.108	1.12	0.267
Rain_t-4	0.00011	0.00007	1.60	0.114
Temp_t-4	-0.0066	0.0132	-0.50	0.617
NDVI_t-5	-0.236	0.097	-2.43	0.017**
Rain_t-5	0.00006	0.00007	0.83	0.411
Temp_t-5	0.021	0.0108	1.99	0.049**
Constant	-0.432	0.361	-1.20	0.234

Table 7. VAR (5) estimation results for agro-environmental variables in east flores (Rainfall equation)

Variable	Coefficient	Std. Error	t-Statistic	p-value
NDVI_t-1	-409.1	200.3	-2.04	0.044**
Rain_t-1	0.232	0.120	1.94	0.055*
Temp_t-2	61.8	23.3	2.66	0.009***
NDVI_t-4	-337.8	196.4	-1.72	0.089*
NDVI_t-5	297.3	176.1	1.69	0.095*
Constant	-1246	656	-1.90	0.060*

Table 8. VAR (5) estimation results for agro-environmental variables in east flores (Temperature equation)

Variable	Coefficient	Std. Error	t-Statistic	p-value
Temp_t-1	0.817	0.101	8.12	<0.001**
Rain_t-2	-0.00145	0.00069	-2.11	0.038**
Temp_t-3	-0.369	0.129	-2.86	0.005***
Temp_t-4	0.240	0.132	1.81	0.073*
Temp_t-5	0.220	0.108	2.03	0.045**
Constant	11.82	3.63	3.26	0.002***

Notes: OLS estimation (equation-by-equation). Lag order = 5 selected by AIC. Sample size = 114 monthly observations. Significance levels: *** p<0.01, ** p<0.05, * p<0.10. Only significant and marginal coefficients are reported for the rainfall and temperature equations due to conference page limits. Evidence of cross-equation error dependence is reported separately in **Table 9**.

Vegetation dynamics exhibit short-run persistence through $(NDVI_{t-1} = 0.374, p = 0.001)$ and a longer-horizon adjustment at $(NDVI_{t-5} = -0.236, p = 0.017)$, while delayed temperature effects emerge at a five-month horizon $(Temp_{t-5} = 0.021, p = 0.049)$. Rainfall and temperature equations similarly display cross-variable lagged effects, reflecting feedback mechanisms typical of agro-environmental systems. In addition to lagged dynamics, the empirical VAR exhibits pronounced contemporaneous dependence across equations. Residual correlations between NDVI and rainfall (−0.494) and between rainfall and temperature (−0.317), together with a highly significant LM test for cross-equation residual correlation $(LM = 40.40, p < 10^{-8})$, indicate that environmental shocks affect multiple system components simultaneously.

Table 9. Evidence of cross-equation error dependence in the empirical VAR (5)

Evidence	Statistic / Value	Interpretation
Residual correlation: NDVI–Rain	−0.493	Moderate negative contemporaneous dependence between NDVI and rainfall residuals
Residual correlation: Rain–Temp	−0.3174	Moderate negative contemporaneous dependence between rainfall and temperature residuals
Residual correlation: NDVI–Temp	0.0996	Weak positive contemporaneous dependence between NDVI and temperature residuals
LM test for cross-equation residual correlation	LM = 40.4016, df = 3, p = 8.76×10^{-9}	Strongly rejects Ho : no cross-equation residual correlation; supports non-diagonal disturbance covariance structure

Despite this dependence, estimation using feasible GLS via SUR yields coefficients identical to equation-by-equation OLS, reflecting the shared regressor structure across equations. This empirical finding reinforces the central message that efficiency gains from GLS are conditional on model structure and covariance estimation rather than automatic. Taken together, the empirical evidence supports the following agro-environmental VAR representation for East Flores:

$$\mathbf{Y}_t = \mathbf{c} + \sum_{i=1}^{5} \mathbf{A}_i \mathbf{Y}_{t-i} + \boldsymbol{\varepsilon}_t, \qquad \boldsymbol{\varepsilon}_t \sim N(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\mathbf{Y}_t = (\mathrm{NDVI}_t, \mathrm{Rain}_t, \mathrm{Temp}t)'$, the coefficient matrices $\mathbf{A}_i$ capture delayed vegetation–climate feedbacks over multiple horizons, and the disturbance covariance matrix $\boldsymbol{\Sigma}$ is non-diagonal, reflecting contemporaneously correlated environmental shocks. This specification provides a coherent modeling framework that accommodates both dynamic interdependence and shared disturbances in agro-environmental systems, while allowing for robust estimation in finite samples. The impulse response and variance decomposition analyses (**Figure 6**) further illustrate the dynamic implications of the estimated model. Rainfall shocks generate delayed and moderate responses in NDVI, while temperature and precipitation shocks contribute increasingly to vegetation variability at longer horizons, underscoring the multivariate nature of agro-environmental dynamics in East Flores.

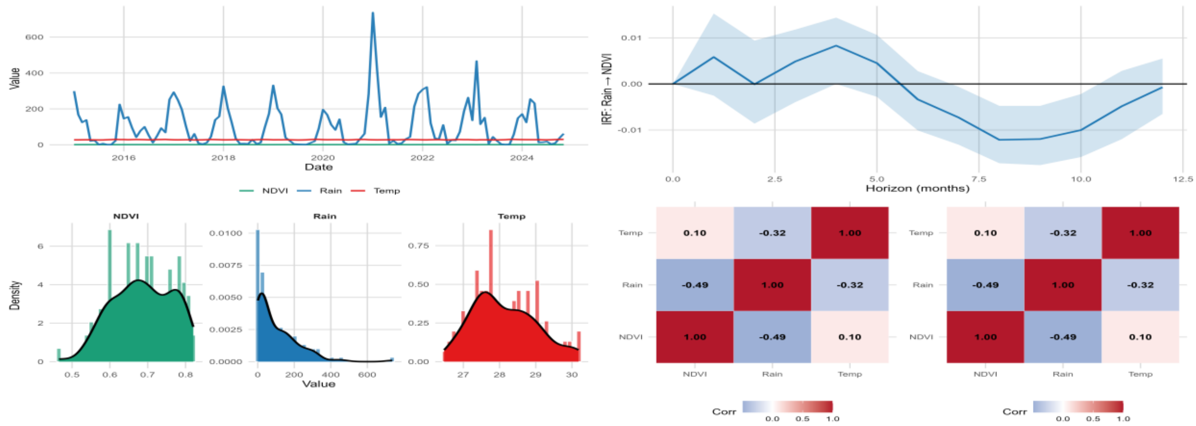

Figure 6. Empirical dynamics, impulse responses, and contemporaneous residual correlations for the VAR(5) model of NDVI, rainfall, and temperature in East Flores

Periods of increased volatility are visible across multiple variables, particularly in recent years, indicating that agro-environmental indicators in East Flores are subject to common environmental shocks rather than purely idiosyncratic fluctuations. This visual evidence is consistent with the presence of contemporaneous dependence among disturbances, as further supported by the residual correlation analysis, and motivates careful consideration of the disturbance covariance structure in multivariate modeling. The empirical results align closely with the structural features highlighted in Scenario 2 of the simulation study, where disturbances are contemporaneously correlated. While the simulation results illustrate that the finite-sample performance of GLS depends critically on the accuracy of covariance estimation, the empirical analysis confirms the practical relevance of correlated disturbances in real agro-environmental data. Taken together, these findings indicate that the simulation framework captures key characteristics of the empirical system, thereby strengthening the methodological contribution of the study without implying that efficiency gains from GLS estimation are automatic in practice.

Discussion: Implications for Real-World Agro-Environmental Analysis

The empirical findings underscore that agro-environmental systems in climate-sensitive regions such as East Flores are characterized by complex dynamic interdependencies and shared environmental disturbances. Climatic variability, together with episodic shocks such as volcanic activity, can generate contemporaneous correlation among disturbances in multivariate time-series models, as evidenced by the residual dependence observed in the empirical VAR analysis. In such settings, equation-by-equation OLS estimation can yield reasonable coefficient estimates, yet it may not fully reflect the underlying dependence structure of environmental shocks. While GLS-type approaches are designed to account for correlated disturbances, the empirical results demonstrate that efficiency gains from GLS are not automatic and depend critically on model structure and covariance estimation quality. The alignment between the simulation framework and the empirical evidence therefore highlights the practical importance of diagnosing disturbance dependence and finite-sample behavior, rather than assuming uniform superiority of any single estimation method, when analyzing real-world agro-environmental systems.

Conclusion

This study demonstrates that agro-environmental dynamics in East Flores are best characterized by a multivariate modeling framework that jointly captures delayed vegetation–climate feedbacks and contemporaneously correlated environmental disturbances. Through a combination of Monte Carlo simulation and empirical VAR analysis, the results show that while correlated shocks are empirically present, efficiency gains from GLS estimation are conditional on covariance estimation quality and model structure rather than automatic. The empirical evidence supports a stable VAR(5) specification with a non-diagonal disturbance covariance matrix as an appropriate modeling representation for NDVI, precipitation, and temperature interactions in disturbance-prone regions. More broadly, the findings emphasize that careful attention to error structures and finite-sample properties is essential when modeling agro-environmental systems, offering practical guidance for applied researchers working with complex, climate-sensitive time-series data.

Recommendations

Based on the findings of this study, several methodological and practical recommendations can be proposed for agro-environmental time-series analysis. First, researchers applying VAR models to agro-environmental systems are encouraged to carefully assess the underlying disturbance structure prior to estimation. Diagnostic checks for contemporaneous correlation and heteroskedasticity should be treated as an integral step in multivariate modeling, particularly in climate-sensitive regions. Second, when evidence of correlated disturbances is present, researchers should consider GLS or feasible GLS as alternative estimation approaches, while recognizing that efficiency gains are conditional rather than automatic. As demonstrated by both the simulation scenarios and the empirical analysis, the performance of GLS depends critically on covariance estimation accuracy, sample size, and model structure, and should therefore be evaluated in conjunction with equation-by-equation OLS rather than assumed to dominate it. Third, simulation-based evaluation is recommended as a complementary tool in applied agro-environmental research. Simulation studies enable researchers to explore estimator behavior under controlled dependence structures and to better understand potential finite-sample distortions before applying multivariate models to real-world data. From a practical perspective, policymakers and practitioners using agro-environmental indicators such as NDVI should be aware that environmental disturbances, including climatic extremes and volcanic activity, may introduce additional variability and dependence in observed data. Explicitly accounting for these features in statistical modeling can improve the robustness and interpretability of analyses used to support agricultural planning and environmental monitoring. Future research may extend the current framework by incorporating additional agro-environmental variables, alternative dependence structures, or spatially disaggregated data to further enhance understanding of vegetation–climate dynamics under environmental variability.

Scientific Ethics Declaration

* This article was presented as an oral presentation at the International Conference on Veterinary, Agriculture and Life Sciences (www.icvals.net) held in Konya/Türkiye on April 02-05, 2026.

* The authors declare that this study does not involve human participants or animal subjects, and therefore ethical approval was not required. The authors take full scientific and legal responsibility for the content of this article.

Conflict of Interest

* The authors declare that they have no conflicts of interest.

Funding

* This research was supported by the Lembaga Pengelola Dana Pendidikan (LPDP), Ministry of Finance of the Republic of Indonesia, under the LPDP Scholarship Program. The funding body had no role in the design of the study; in the collection, analysis, or interpretation of data; or in the writing of the manuscript.

Acknowledgements or Notes

* This article was presented as an oral/poster presentation at the International Conference on Veterinary, Agriculture and Life Sciences (www.icvals.net), held in Konya, Türkiye, on April 2–5, 2026.

* The author would like to express sincere gratitude to the LPDP Scholarship Program for financial support, and to all institutions and colleagues who provided valuable insights and support during the preparation of this study.

References

Akpan, E., & Moffat, I. (2018). Modeling the autocorrelated errors in time series regression: A generalized least squares approach. *Journal of Advances in Mathematics and Computer Science*, *26*(4), 1–15. https://doi.org/10.9734/jamcs/2018/39949

Arai, Y., Otsu, T., & Xu, M. (2024). GLS under monotone heteroskedasticity. *Journal of Econometrics*, *246*(1–2), 105899. https://doi.org/10.1016/j.jeconom.2024.105899

Bai, J., Choi, S. H., & Liao, Y. (2021). Feasible generalized least squares for panel data with cross-sectional and serial correlations. *Empirical Economics*, *60*(1), 309–326. https://doi.org/10.1007/s00181-020-01977-2

Cordoni, F., Dorémus, N., & Moneta, A. (2024). Identification of vector autoregressive models with nonlinear contemporaneous structure. *Journal of Economic Dynamics and Control*, *162*, 104852. https://doi.org/10.1016/j.jedc.2024.104852

Feng, S., Zhang, Z., Zhao, S., Guo, X., Zhu, W., & Das, P. (2023). Time lag effect of vegetation response to seasonal precipitation in the Mara River Basin. *Ecological Processes*, *12*(1), 1–13. https://doi.org/10.1186/s13717-023-00461-w

Güler, S., & Turgut, B. (2024). Monitoring Türkiye's vegetation cover with NDVI: Terrestrial and temporal perspectives. *SilvaWorld*, *3*(2), 87–98. https://doi.org/10.61326/silvaworld.v3i2.285

Guo, J., Wang, K., Wang, T., Bai, N., Zhang, H., Cao, Y., & Liu, H. (2022). Spatiotemporal variation of vegetation NDVI and its climatic driving forces in global land surface. *Polish Journal of Environmental Studies*, *31*(4), 3541–3549. https://doi.org/10.15244/pjoes/147194

Kadhim Tayyeh, H., & Mohammed, R. (2023). Analysis of NASA POWER reanalysis products to predict temperature and precipitation in Euphrates River basin. *Journal of Hydrology*, *619*, 129327. https://doi.org/10.1016/j.jhydrol.2023.129327

Li, A., Yin, S., Li, N., & Shi, C. (2025). Comprehensive analysis of the driving forces behind NDVI variability in China under climate change conditions and future scenario projections. *Atmosphere*, *16*(6), 738. https://doi.org/10.3390/atmos16060738

Liu, M., Zhai, H., Zhang, X., Dong, X., Hu, J., Ma, J., & Sun, W. (2024). Time-lag and accumulation responses of vegetation growth to average and extreme precipitation and temperature events in China between 2001 and 2020. *Science of The Total Environment*, *945*, 174084. https://doi.org/10.1016/j.scitotenv.2024.174084

Lütkepohl, H. (2005). *New introduction to multiple time series analysis*. Berlin, Germany: Springer-Verlag.

Martinez, L., Roberts, A., & Oliveira, P. (2021). Impact of environmental shocks on vegetation time series in tropical regions. *ISPRS Journal of Photogrammetry and Remote Sensing*, *175*, 138–152. https://doi.org/10.1016/j.isprsjprs.2021.03.003

Nguyen, T., Zhang, F., & Chen, Y. (2022). Assessing vegetation–climate interactions using VAR and cointegration models. *Remote Sensing of Environment*, *277*, 113058. https://doi.org/10.1016/j.rse.2022.113058

Ogutu, J. O., Bartzke, G. S., Mukhopadhyay, S., Dublin, H. T., Senteu, J. S., Gikungu, D., Obara, I., & Piepho, H.-P. (2024). Trends and cycles in rainfall, temperature, NDVI, IOD and SOI in the Mara-Serengeti: Insights for biodiversity conservation. *PLOS Climate*, *3*(10), e0000388. https://doi.org/10.1371/journal.pclm.0000388

Qin, J., Ma, M., Shi, J., Ma, S., Wu, B., & Su, X. (2023). The time-lag effect of climate factors on the forest enhanced vegetation index for subtropical humid areas in China. *International Journal of Environmental Research and Public Health*, *20*(1), 799. https://doi.org/10.3390/ijerph20010799

Rodrigues, G. C., & Braga, R. P. (2021). Evaluation of NASA POWER reanalysis products to estimate daily weather variables in a hot summer Mediterranean climate. *Agronomy*, *11*(6), 1207. https://doi.org/10.3390/agronomy11061207

Sims, C. A. (1980). Macroeconomics and reality. *Econometrica*, *48*(1), 1–48. https://doi.org/10.2307/1912017

Sun, M., Gong, A., Zhao, X., Liu, N., Si, L., & Zhao, S. (2023). Reconstruction of a monthly 1 km NDVI time series product in China using Random Forest methodology. *Remote Sensing*, *15*(13), 3353. https://doi.org/10.3390/rs15133353

Wu, Y., Zhang, X., Fu, Y., Hao, F., & Yin, G. (2020). Response of vegetation to changes in temperature and precipitation at a semi-arid area of northern China based on multi-statistical methods. *Forests*, *11*(3), 340. https://doi.org/10.3390/f11030340

Xiong, C., Ma, H., Liang, S., He, T., Zhang, Y., Zhang, G., & Xu, J. (2023). Improved global 250 m 8-day NDVI and EVI products from 2000–2021 using the LSTM model. *Scientific Data*, *10*(1), 1–14. https://doi.org/10.1038/s41597-023-02695-x

Author(s) Information

Agnes Ona Bliti Puka
Sepuluh Nopember Institute of Technology, Department of Statistics, Faculty of Science and Data Analytics, Kampus ITS-Sukolilo, Surabaya 60111, Indonesia
Institut Keguruan dan Teknologi Larantuka, Department of Mathematics Education, Faculty of Teacher Training and Education, Jl. Ki Hajar Dewantara, Kampus IKTL Waibalun, Larantuka, 86212, Indonesia
ORCID iD: https://orcid.org/0009-0001-9779-8885

Kartika Fithriasari
Sepuluh Nopember Institute of Technology, Department of Statistics, Faculty of Science and Data Analytics, Kampus ITS-Sukolilo, Surabaya 60111, Indonesia
*Contact e-mail: *kartika_f@its.ac.id*
ORCID iD: https://orcid.org/0000-0003-4543-4884

Dedy Dwi Prastyo
Sepuluh Nopember Institute of Technology, Department of Statistics, Faculty of Science and Data Analytics, Kampus ITS-Sukolilo, Surabaya 60111, Indonesia
ORCID iD: https://orcid.org/0000-0003-1194-769X

*Corresponding author(s) contact email

To cite this article:

Puka, A.O.B., Fithriasari, K., & Prastyo, D.D. (2026). Modeling agro-environmental dynamics with VAR under climate and volcanic variability: A simulation-based OLS–GLS comparison from east flores, Indonesia. *The Eurasia Proceedings of Health, Environment and Life Sciences (EPHELS), 21,* 18-32.