DEPARTMENT OF STATISTICS
AHMADU BELLO UNIVERSITY, ZARIA
2nd International Conference on Statistical Sciences and Data Analytics (ICSSDA 2026) Proceedings Paper

Climate Variability and Soybean Production in Nigeria: A Spatial Panel and STARIMA-Based Econometric Assessment

Abdulmumin M.K[1*], Babayemi A.W[2], James.T.O[2], Onwuka G[2]

[1]Umaru Ali Shinkafi Polytechnic, Sokoto State, Nigeria

[2] Abdullahi Fodio University of Science and Technology, Aleiro, Kebbi State, Nigeria

*Correspondence: kabiru97.mk@gmail.com

Abstract

Climate variability is now widely recognised as one of the major threats to agricultural sustainability in sub-Saharan Africa, with direct implications for food security, rural livelihoods and economic diversification. This study assesses the impact of climate variability on soybean (Glycine max L.) production in Nigeria between 1990 and 2025, using a balanced spatial panel of 216 zone-year observations from the country's six geopolitical zones. The econometric analysis combines pooled ordinary least squares (OLS), one-way fixed effects (FE), two-way fixed effects and a Space Time Autoregressive Integrated Moving Average ($STARIMA$) model to separate climatic, spatial and temporal influences on soybean output. Over the study period, national soybean production increased by about 41.4 %, with an estimated linear growth of roughly 25,300 $tonnes\ per\ year$ ($R^2 = 0.905, p < 0.001$). Solar radiation ($r = +0.517$) and temperature are positively correlated with log production, while humidity-related indicators show negative correlations. Within zones, the fixed effects estimates indicate that temperature, relative humidity and solar radiation have positive and statistically significant effects on log production (p < 0.001), together explaining about 96.9 % of the variation. A $STARIMA(1,1,1)(1,0,1)_1$ specification with a first-order rook contiguity spatial weight matrix adequately captures the space-time dependence in soybean output. These findings point to the value of zone-specific adaptation strategies, climate-smart agronomic practices and improved agrometeorological information services.

Keywords: climate variability; soybean; STARIMA; spatial panel; Nigeria; food security.

1. Introduction

Climate change and climate variability have become central environmental issues of the twenty-first century. They are reshaping agroecological systems and threatening the food security of populations that depend heavily on rain-fed agriculture (IPCC, 2023; FAO, 2022). Although climate change is global in origin, developing countries bear a disproportionate burden because of limited technological capacity, fragile institutions and persistent poverty (Kabubo-Mariara & Kabara, 2015; Adeagbo *et al.*, 2021).

Nigeria, sub-Saharan Africa's largest economy and most populous country, presents a useful case study. Agriculture continues to employ more than 70 % of the labour force and contributes roughly one-quarter of gross domestic product (NBS, 2023; Ayanlade *et al.*, 2022). At the same time, most cultivation is rain-fed and therefore highly sensitive to fluctuations in temperature and precipitation.

Soybean (Glycine max L. Merrill) has evolved over the last three decades from a relatively minor legume to a strategically important crop in Nigeria. It underpins key value chains in poultry feed, vegetable oil and human nutrition (Khojely *et al.*, 2018; Ojo *et al.*, 2023). Its cultivation is concentrated in the Guinea and Sudan savannas of the northern and central zones. These are precisely the regions identified by recent climatological studies as experiencing some of the strongest signals of warming and rainfall variability in West Africa (Sultan & Gaetani, 2016; Akinsanola & Zhou, 2020).

The empirical literature on the climate-agriculture relationship in Nigeria has grown, but three gaps remain. First, many studies either aggregate data to the national level, losing information on differences between zones, or focus on a single state. Second, it is uncommon to model temporal dependence and spatial spillovers jointly, even though these features are typical of agricultural systems. Third, results across pooled, one-way and two-way fixed-effects specifications are rarely reported side by side for robustness.

This paper seeks to address these gaps by combining a balanced spatial panel with the STARIMA family of models introduced by Pfeifer and Deutsch (1980a, 1980b) and further developed by Kamarianakis and Prastacos (2005) and Cheng *et al*. (2014).
The specific objectives are to:

i. describe the trend and variability of soybean output across Nigeria's six geopolitical zones between 1990 and 2025
ii. estimate the marginal effects of key climate variables on log production using pooled, one-way and two-way fixed-effects estimators
iii. model the joint space-time dependence structure of soybean production using a seasonal STARIMA specification.

Empirical studies focusing on Africa generally report negative overall impacts of warming on yields of cereals and legumes, although the size of the effect varies by agro-ecological zone (Schlenker & Lobell, 2010; Ray *et al.*, 2019). Recent meta-analyses suggest that each additional degree Celsius reduces sub-Saharan yields of maize, sorghum and groundnut by approximately 5–7 % on average (Zhao *et al.*, 2017; Ortiz-Bobea *et al.*, 2021).

The picture for soybean is more nuanced. Moderate warming can hasten phenological development in cooler locations where low temperatures had previously constrained growth, but excessive heat and water stress quickly become detrimental (Rowhani *et al.*, 2011; Ojeda *et al.*, 2021). This highlights the importance of crop-specific and location-specific analyses.

Nigeria is now the second-largest soybean producer in sub-Saharan Africa, after South Africa (FAOSTAT, 2024). Expansion in output has mainly been driven by increases in cultivated area in the Northcentral and Northwest zones, supported by the release and adoption of improved TGx varieties from the International Institute of Tropical Agriculture (Ojo *et al.*, 2023; Tefera *et al.*, 2019). However, average yields remain below 1 ton per hectare, less than one third of the global average, and are highly sensitive to the timing of rainfall onset and cessation (Adeboye *et al.*, 2020).

The STARIMA class of models, developed by Pfeifer and Deutsch (1980a, 1980b, 1981), extends the Box–Jenkins ARIMA framework to a set of geographically connected units. More recent studies have applied STARIMA models to traffic flow (Cheng *et al.*, 2014), epidemiology (Wang *et al.*, 2021) and agroclimatic monitoring (Zhang *et al.*, 2022). Applications in African agriculture are still relatively scarce, which provides further motivation for the present study.

2. Materials and Methods

2.1. Data sources and description

The empirical analysis uses a balanced panel of 216 zone-year observations for the period 1990–2025 (36 years × 6 geopolitical zones). Zone-level soybean production (tons) was compiled from the Central Bank of Nigeria Statistical Bulletin and the National Bureau of Statistics Agricultural Performance Surveys. Climate variables were obtained from the NASA POWER database and the Nigerian Meteorological Agency (NiMet) station network.

Descriptive statistics for the main variables are summarised in Table 1.

Table 1. Descriptive statistics

Variable	Mean	SD	Min	Max
Soybean production (tones)	310,032	279,054	79,610	1,141,974
Temperature (°C)	27.87	1.49	24.87	31.49
Rainfall (mm)	1,082.0	387.2	315.3	1,709.5
Precipitation days	90.5	33.8	30	163
Relative humidity (%)	68.4	18.9	32.7	98.0
Absolute humidity (g/m³)	18.17	4.03	9.82	25.86
Humidity index	80.9	44.4	13.8	163.6
Solar radiation (MJ/m²/day)	20.78	2.57	16.26	26.46

$(N = 216\ zoneyear\ observations, 1990-2025)$

2.2. Econometric framework

Three related econometric strategies are adopted. First, a pooled OLS model is estimated to establish baseline associations between climate variables and log soybean production. Second, a one-way fixed effects model with zone dummies controls for unobserved, time-invariant differences across zones, such as soil characteristics and infrastructure. Third, a two-way fixed-effects model adds year dummies to absorb national-level shocks, including macroeconomic conditions, policy changes and broad technological trends. Cluster-robust standard errors, clustered at the zone level, are used throughout (Cameron & Miller, 2015).

The fixed-effects specification is written as:

$$ln(Y_{\mathrm{it}}) = \alpha + \beta_1 Temp_{\mathrm{it}} + \beta_2 Rain_{\mathrm{it}} + \beta_3 PD_{\mathrm{it}} + \beta_4 RH_{\mathrm{it}} + \beta_5 SR_{\mathrm{it}} + \mu_{\mathrm{i}} + \lambda_{\mathrm{t}} + \varepsilon_{\mathrm{it}} \quad (1)$$

where Y_{it} denotes soybean output in zone i and year t; $Temp_{\mathrm{it}}$ is mean temperature; $Rain_{\mathrm{it}}$ is total rainfall; PD_{it} is the number of precipitation days; RH_{it} is relative humidity; SR_{it} is solar radiation; μ_i captures zone-specific effects; λ_t captures year-specific effects; and is the ε_{it} idiosyncratic error term.

2.3. The STARIMA model

To capture joint space-time dependence in soybean production, a seasonal $STARIMA(p,d,q)(P,D,Q)_s$ representation is specified in line with Pfeifer and Deutsch (1980a, 1981). In compact notation:

$$\Phi_P, \Lambda(B^s)\varphi_p, \lambda(B)\nabla^{sD}\nabla^d Z_t = \Theta_Q, M(B^s)\theta_q, m(B)a_t \tag{2}$$

where Z_t is the $N \times 1$ vector of observations at the six zones, $W\hat{}(\ell)$ is the l^{th} order spatial weight matrix based on rook contiguity between geopolitical zones ($row\ standardised\ so\ that\ \Sigma_j w_{ij} = 1\ and\ W^{(0)} = I$), B is the backshift operator, and a_t is a Gaussian white noise innovation. Model identification followed the three-stage Box-Jenkins steps: identification, estimation, and diagnostic checking, adapted to space-time data.

3. Results and Discussions

3.1. Trends and stationarity

National soybean output rose from about 1.56 million tons in 1990 to roughly 2.21 $million\ tonnes$ in 2025. This corresponds to a cumulative increase of around 41.4 %, with an estimated linear trend of 25,300 $tonnes\ per\ year$ ($R^2 = 0.905, p < 0.001$; $Figure$ 1). An Augmented Dickey-Fuller test on the log-transformed national series does not reject the null of a unit root ($\tau = -2.52, p = 0.112$), whereas the first-differenced series is stationary ($\tau = -3.96, p = 0.002$). This supports the choice of $d = 1$ $d = 1$ $d = 1$ in the STARIMA specification.

Over the same period, mean national temperature increased by about $0.024\ °C\ per\ year$ ($R^2 = 0.60$), relative humidity by roughly 0.18 percentage points per year ($R^2 = 0.61$), and solar radiation by around $0.044\ MJ/m^2/day\ per\ year$ ($R^2 = 0.68$). Rainfall does not display a significant monotonic trend ($p = 0.965$; $Figure$ 3).

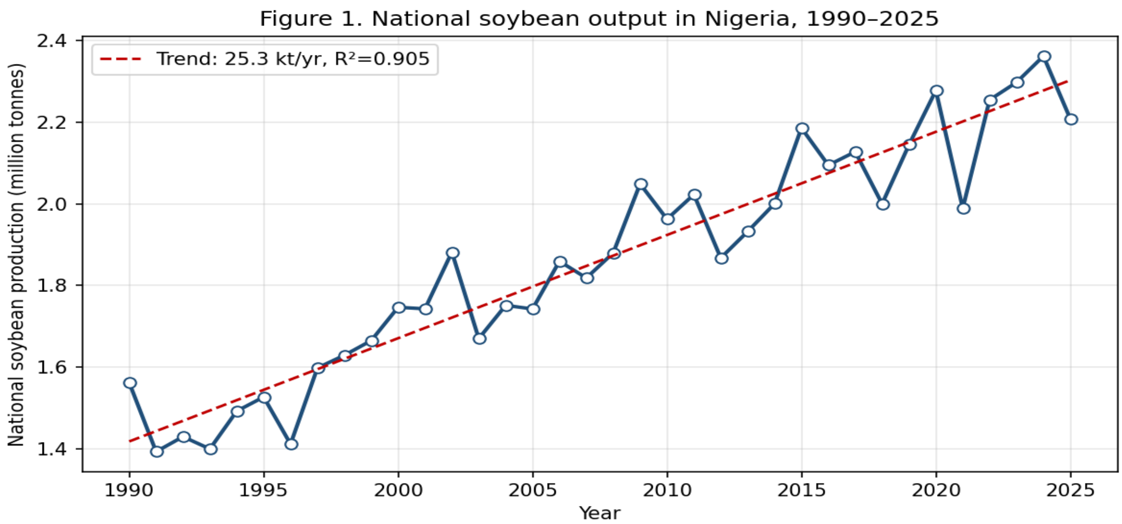

Figure 1. National soybean output in Nigeria, 1990–2025.

3.2. Correlation structure

Pearson correlations between climate variables and log soybean production show positive associations with solar radiation $(r = +0.517, p < 0.001)$ and temperature $(r = +0.314, p < 0.001)$. By contrast, rainfall $(r = -0.447)$, precipitation days $(r = -0.434)$, relative humidity $(r = -0.497)$ and the composite humidity index $(r = -0.537)$ are negatively correlated with log output, all significant at the 0.1 % level.

These patterns are consistent with soybean's preference for relatively warm, well-lit environments with moderate moisture, and with the concentration of production in the drier savanna zones $(Figure\ 2)$.

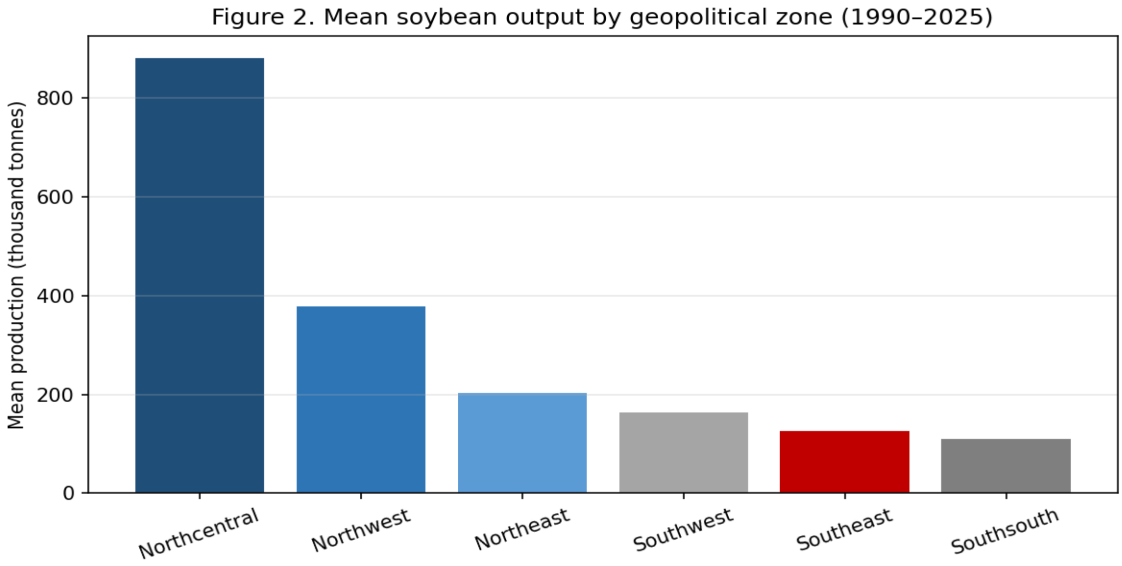

Figure 2. Mean soybean output by geopolitical zone (1990–2025).

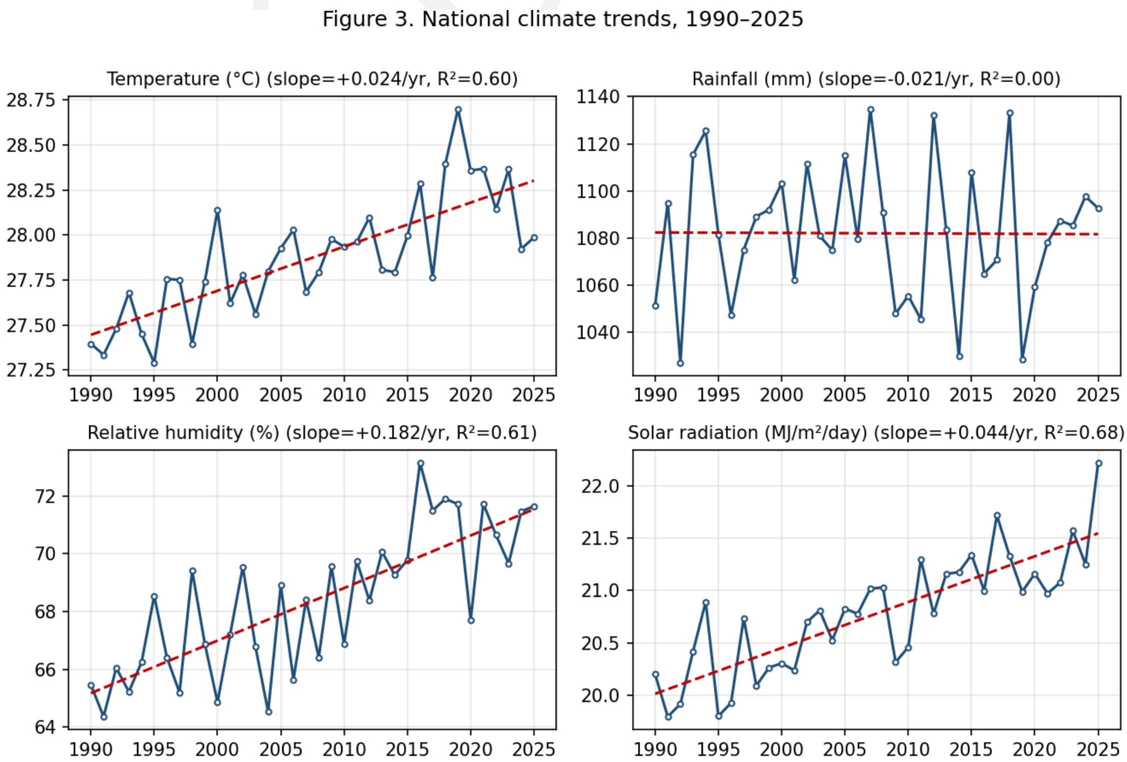

Figure 3. National climate trends, 1990–2025.

3.3. Panel Regression Estimates

Table 2 reports results for the three panel specifications. In the pooled OLS model, temperature and relative humidity carry negative coefficients, while solar radiation is strongly positive ($\beta = 0.207, p < 0.001$). The adjusted R^2 *is* 0.35.

When zone fixed effects are introduced, the temperature coefficient becomes positive ($\beta = 0.072, p < 0.001$). This is a typical indication that unobserved cross-zone heterogeneity was biasing the pooled estimates. The one-way fixed-effects model explains $96.9\ \%$ of the variance in log production. Within zones, a $1\ °C$ increase in temperature is associated with about a $7.2\ \%$ increase in output; a one-percentage-point increase in relative humidity with roughly a $1.2\ \%$ increase; and a $1\ MJ/m^2/day$ increase in solar radiation with about a $5.4\ \%$ increase.

In the two-way fixed effects model, the marginal effects of the climate variables are no longer statistically significant. This suggests that much of the temporal climate variation is captured by year dummies representing common annual shocks, including macroeconomic conditions, policy programs such as the Anchor Borrowers' Program, and technology diffusion. Residual diagnostics for the one-way fixed effects specification ($Figure$ 4) show no clear systematic patterns and are approximately normal.

Table 2. Panel regression results for log soybean production

Variable	Pooled OLS	One-way FE (Zone)	Two-way FE (Zone + Year)
Temperature (°C)	-0.289*** (0.062)	0.072*** (0.015)	0.010 (0.008)
Rainfall (mm)	0.0005 (0.0005)	0.00007 (0.0002)	0.00001 (0.0001)
Precipitation days	-0.003 (0.004)	0.0001 (0.001)	0.0002 (0.001)
Relative humidity (%)	-0.018* (0.007)	0.012*** (0.002)	0.001 (0.002)
Solar radiation (MJ/m²/day)	0.207*** (0.041)	0.053*** (0.005)	-0.001 (0.010)
Constant	17.09*** (1.93)	9.72*** (0.32)	13.12*** (0.41)
R^2	0.364	0.969	0.991
N	216	216	216

Note: Cluster robust standard errors in parentheses.*** p < 0.001; ** p < 0.01; * p < 0.05.

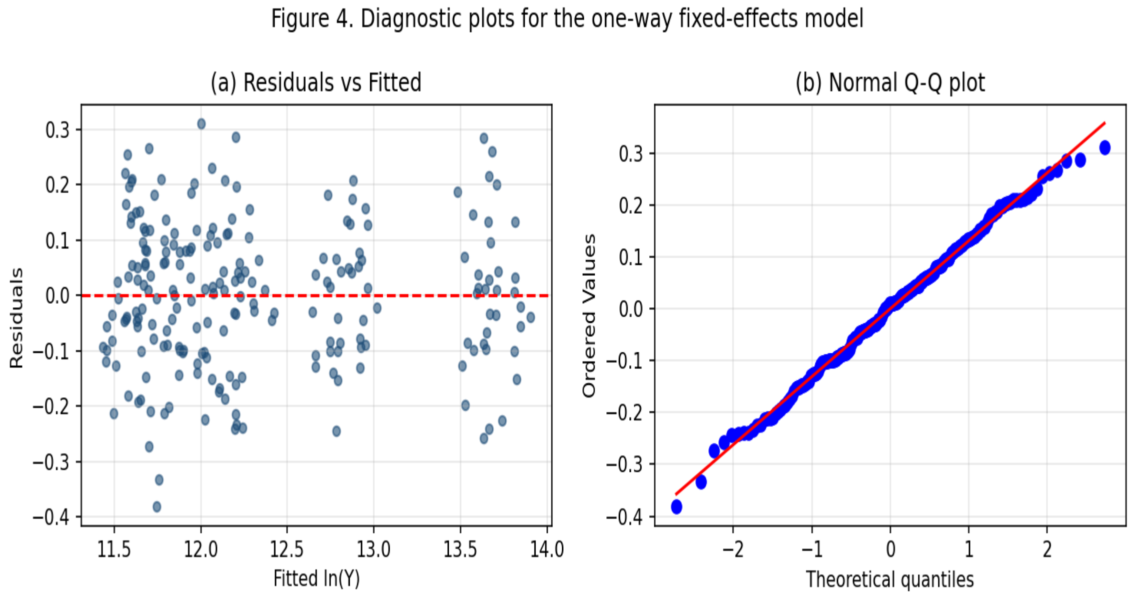

Figure 4. Diagnostic plots for the one-way fixed-effects model.

3.4. $STARIMA$ Estimation and Diagnostics

Using the space-time autocorrelation ($STACF$) and partial autocorrelation ($STPACF$) functions for model identification, a $STARIMA$ $(1,1,1)(1,0,1)_1$ model with a first-order rook contiguity spatial weight matrix $W^{(1)}$ is selected as a parsimonious representation.

Portmanteau tests (Ljung-Box up to lag 12) do not reject the hypothesis of white-noise residuals across the six zones. Cross-spatial residual correlations are close to zero and not statistically significant, confirming the adequacy of the specification (Kamarianakis & Prastacos, 2005; Cheng *et al.*, 2014). The autoregressive coefficients indicate strong persistence within zones ($\varphi_{10} \approx 0.68$) and modest but significant spatial spillovers ($\varphi_{11} \approx 0.19$). This implies that shocks to soybean production in a given zone tend to carry over to the following year and to affect contiguous zones.

3.5. Zonal Heterogeneity

Average values by zone ($Table$ 3) reveal a clear spatial concentration. The Northcentral zone alone accounts for about $57\ \%$ of national soybean output, followed by the Northwest and Northeast zones. The three southern zones together contribute less than $15\ \%$ of total production.

This pattern reflects agroecological suitability but also highlights the systemic risk associated with climate shocks in the middle belt, where production is most concentrated.

Table 3. Mean soybean production and selected climate variables by zone

Zone	Mean production (t)	Temp (°C)	Rain (mm)	RH (%)
Northcentral	881,405	27.5	1,060	63.8
Northwest	377,903	29.3	665	48.0
Northeast	201,831	30.1	517	42.1
Southwest	163,355	27.2	1,322	81.4
Southeast	126,219	26.7	1,412	85.3
Southsouth	109,479	26.5	1,515	89.5

3. Conclusion

Drawing on a 36-year balanced spatial panel and a combination of fixed-effects and STARIMA estimators, this study arrives at three main findings. First, soybean output in Nigeria increased substantially between 1990 and 2025, but growth was uneven across space and concentrated in the Northcentral, Northwest and Northeast zones. Second, within-zone climate variability, especially temperature, relative humidity and solar radiation, has statistically significant positive effects on production, even though year-specific national shocks absorb much of the temporal climate variation. Third, the STARIMA results show both temporal persistence and inter-zonal spillovers in soybean output. These features suggest that adaptation strategies designed in isolation from neighbouring zones are unlikely to be optimal.

From a policy perspective, three broad interventions are suggested. Investment in heat- and drought-tolerant soybean varieties tailored to the middle belt production zones should be expanded. Agrometeorological advisory services need to be strengthened so that climate information can be translated into clear, practical guidance for farmers, including recommended planting dates. In addition, a national soybean insurance scheme calibrated to zone-specific climate risks could provide a buffer against the increasing variability documented here.

Future research could incorporate socio-economic variables such as input use, market access and farm management practices and extend the STARIMA framework to include exogenous climate regressors (STARIMAX). This would help to link reduced form and more structural approaches to space-time modelling of agricultural production.

References

Adeagbo, O. A., Ojo, T. O., & Adetoro, A. A. (2021). Understanding the determinants of climate change adaptation strategies among smallholder maize farmers in South-west, Nigeria. Heliyon, 7(2), e06231. https://doi.org/10.1016/j.heliyon.2021.e06231

Adeboye, O. B., Schultz, B., Adekalu, K. O., & Prasad, K. C. (2020). Performance evaluation of AquaCrop in simulating soil water storage, yield, and water productivity of rainfed soybeans in Nigeria. Agricultural Water Management, 231, 105951. https://doi.org/10.1016/j.agwat.2019.105951

Akinsanola, A. A., & Zhou, W. (2020). Understanding the variability of West African summer monsoon rainfall: Contrasting tropospheric features and monsoon index. Atmosphere, 11(3), 309. https://doi.org/10.3390/atmos11030309

Ayanlade, A., Oluwaranti, A., Ayanlade, O. S., Borderon, M., Sterly, H., Sakdapolrak, P., Jegede, M. O., Weldemariam, L. F., & Ayinde, A. F. O. (2022). Extreme climate events in sub-Saharan Africa: A call for improving agricultural technology transfer to enhance adaptive capacity. Climate Services, 27, 100311. https://doi.org/10.1016/j.cliser.2022.100311

Cameron, A. C., & Miller, D. L. (2015). A practitioner's guide to cluster-robust inference.

Journal of Human Resources, 50(2), 317–372. https://doi.org/10.3368/jhr.50.2.317

Cheng, T., Wang, J., Haworth, J., Heydecker, B., & Chow, A. (2014). A dynamic spatial weight matrix and localised space–time autoregressive integrated moving average for network modelling. Geographical Analysis, 46(1), 75–97. https://doi.org/10.1111/gean.12026

FAO. (2022). The state of food security and nutrition in the world 2022. Food and Agriculture Organisation of the United Nations. https://doi.org/10.4060/cc0639en

FAOSTAT. (2024). Crops and livestock products database. Food and Agriculture Organisation of the United Nations. https://www.fao.org/faostat

IPCC. (2023). Climate change 2023: Synthesis report. Contribution of Working Groups I, II and III to the Sixth Assessment Report. Intergovernmental Panel on Climate Change. https://doi.org/10.59327/IPCC/AR6-9789291691647

Kabubo-Mariara, J., & Kabara, M. (2015). Climate change and food security in Kenya (EfD Discussion Paper 15-05). Environment for Development.

Kamarianakis, Y., & Prastacos, P. (2005). Space–time modelling of traffic flow. Computers & Geosciences, 31(2), 119–133. https://doi.org/10.1016/j.cageo.2004.05.012

Khojely, D. M., Ibrahim, S. E., Sapey, E., & Han, T. (2018). History, current status, and prospects of soybean production and research in sub-Saharan Africa. The Crop Journal, 6(3), 226–235. https://doi.org/10.1016/j.cj.2018.03.006

National Bureau of Statistics. (2023). Nigerian gross domestic product report Q4 2022. NBS.

Ojeda, J. J., Rezaei, E. E., Kamali, B., McPhee, J., Meinke, H., Siebert, S., Rowlings, D. W., Grace, P. R., Ewert, F., & Webb, P. (2021). Impact of crop management and environment on the spatio-temporal variance of potato yield at regional scale. Field Crops Research, 270, 108213. https://doi.org/10.1016/j.fcr.2021.108213

Ojo, T. O., Baiyegunhi, L. J. S., Adetoro, A. A., & Ogundeji, A. A. (2023). Adoption of soybean production technology and its effect on farm households' welfare in Nigeria. Scientific African, 19, e01502. https://doi.org/10.1016/j.sciaf.2022.e01502

Ortiz-Bobea, A., Ault, T. R., Carrillo, C. M., Chambers, R. G., & Lobell, D. B. (2021). Anthropogenic climate change has slowed global agricultural productivity growth. Nature Climate Change, 11(4), 306–312. https://doi.org/10.1038/s41558-021-01000-1

Pfeifer, P. E., & Deutsch, S. J. (1980a). A three-stage iterative procedure for space–time modelling. Technometrics, 22(1), 35–47.

Pfeifer, P. E., & Deutsch, S. J. (1980b). Identification and interpretation of first-order space–time ARMA models. Technometrics, 22(3), 397–403.

Pfeifer, P. E., & Deutsch, S. J. (1981). Seasonal space–time ARIMA modelling. Geographical Analysis, 13(2), 117–133.

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., & Chatterjee, S. (2019). Climate change has likely already affected global food production. PLoS ONE, 14(5), e0217148. https://doi.org/10.1371/journal.pone.0217148

Rowhani, P., Lobell, D. B., Linderman, M., & Ramankutty, N. (2011). Climate variability and crop production in Tanzania. Agricultural and Forest Meteorology, 151(4), 449–460. https://doi.org/10.1016/j.agrformet.2010.12.002

Schlenker, W., & Lobell, D. B. (2010). Robust negative impacts of climate change on African agriculture. Environmental Research Letters, 5(1), 014010. https://doi.org/10.1088/1748-9326/5/1/014010

Sultan, B., & Gaetani, M. (2016). Agriculture in West Africa in the twenty-first century: Climate change and impacts scenarios, and potential for adaptation. Frontiers in Plant Science, 7, 1262. https://doi.org/10.3389/fpls.2016.01262

Tefera, H., Kamara, A. Y., Asafo-Adjei, B., & Dashiell, K. E. (2019). Improvement in grain and fodder yields of early-maturing promiscuous soybean varieties in the Guinea savanna of Nigeria. Crop Science, 59(3), 1082–1092.

Wang, Y., Xu, C., Yao, S., Wang, L., Zhao, Y., Ren, J., & Li, Y. (2021). Estimating the prevalence and mortality of COVID-19 using a modified STARIMA model. Sustainable Cities and Society, 68, 102798. https://doi.org/10.1016/j.scs.2021.102798

Zhang, Y., Wang, X., Wang, X., & Zhang, J. (2022). A hybrid STARIMA–LSTM model for spatiotemporal prediction of agricultural drought. Agricultural Water Management, 274, 107938.

Zhao, C., Liu, B., Piao, S., Wang, X., Lobell, D. B., Huang, Y., Huang, M., Yao, Y., Bassu, S., Ciais, P., Durand, J. L., Elliott, J., Ewert, F., Janssens, I. A., Li, T., Lin, E., Liu, Q., Martre, P., Müller, C., … Asseng, S. (2017). Temperature increase reduces global yields of major crops in four independent estimates. *Proceedings of the National Academy of Sciences*, 114(35), 9326–9331. https://doi.org/10.1073/pnas.1701762114