

RESEARCH

Open Access



Optimizing spatio-temporal correlation structures for modeling food security in Africa: a simulation-based investigation

Adusei Bofa^{1*} and Temesgen Zewotir¹

*Correspondence:
221119873@stu.ukzn.ac.za; bofa.adusei@gcuc.edu.gh

¹ School of Mathematics, Statistics, and Computer Science, University of KwaZulu Natal, Oliver Tambo Building, Westville Campus, Durban, South Africa

Abstract

This study investigates the impact of spatio-temporal correlation using four spatio-temporal models: Spatio-Temporal Poisson Linear Trend Model (SPLTM), Poisson Temporal Model (TMS), Spatio-Temporal Poisson Anova Model (SPAM), and Spatio-Temporal Poisson Separable Model (STSM) concerning food security and nutrition in Africa. Evaluating model goodness of fit using the Watanabe Akaike Information Criterion (WAIC) and assessing bias through root mean square error and mean absolute error values revealed a consistent monotonic pattern. SPLTM consistently demonstrates a propensity for overestimating food security, while TMS exhibits a diverse bias profile, shifting between overestimation and underestimation based on varying correlation settings. SPAM emerges as a beacon of reliability, showcasing minimal bias and WAIC across diverse scenarios, while STSM consistently underestimates food security, particularly in regions marked by low to moderate spatio-temporal correlation. SPAM consistently outperforms other models, making it a top choice for modeling food security and nutrition dynamics in Africa. This research highlights the impact of spatial and temporal correlations on food security and nutrition patterns and provides guidance for model selection and refinement. Researchers are encouraged to meticulously evaluate the biases and goodness of fit characteristics of models, ensuring their alignment with the specific attributes of their data and research goals. This knowledge empowers researchers to select models that offer reliability and consistency, enhancing the applicability of their findings.

Keywords: Bayesian poisson model, Markov chain monte carlo(MCMC), Matrix plot, Mean absolute error, Root mean square error, Watanabe akaike information criterion

Introduction

Food security is a critical concern, particularly in the context of Africa, where numerous countries face persistent challenges in ensuring adequate access to safe, nutritious, and affordable food for their populations. Understanding the complex dynamics of food security is essential for policymakers, researchers, and practitioners to develop effective interventions and strategies to address this pressing issue.



© The Author(s) 2024. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>. The Creative Commons Public Domain Dedication waiver (<http://creativecommons.org/publicdomain/zero/1.0/>) applies to the data made available in this article, unless otherwise stated in a credit line to the data.

Traditionally, food security analysis has focused on temporal aspects, examining trends and changes in food availability, access, and utilization over time [1]. Nicholson, Stephens [1] reiterated the overemphasis on availability and accessibility indicators. In their study, they put forth four recommendations: avoiding confusion between "food availability" and "food security," consolidating food access indicators, assessing stable outcomes for food security indicators, and establishing empirical data linking results from agricultural systems models to outcomes related to food access. This emphasizes the necessity of including a broader spectrum of variables when evaluating food security and nutrition, particularly within the African context. However, it is increasingly recognized that food security is not solely determined by temporal factors but is also influenced by spatial dynamics [2]. Cooper, Brown [3] conducted a review of factors associated with food security, analyzing 16,152 abstracts from 3297 publications spanning the years 1975 to 2018. The review revealed that a majority of the studied publications focused solely on spatial effects rather than considering spatio-temporal impacts. This highlights the necessity for exploring spatio-temporal modeling.

To capture spatial–temporal variations and better comprehend the intricacies of food security in Africa, there is a growing need for spatio-temporal modeling approaches. Spatio-temporal models integrate both the spatial and temporal dimensions of data, enabling a more comprehensive understanding of the underlying processes and patterns that shape food security dynamics [4]. Spatio-temporal modeling offers several advantages as it explores spatial dependencies, assesses temporal dependencies, and captures persistence, trends, and cyclical patterns in food security outcomes. Moreover, it provides a framework to investigate interactions between spatial and temporal factors, recognizing complex feedback loops and dynamics within food systems [5].

Over the years, Bayesian-based spatio-temporal modeling has emerged as a significant area of interest due to its versatile applications across various disciplines, including ecology and epidemiology [6]. Aswi, Cramb [7] conducted a substantive case study on Dengue Incidence, exploring a range of Bayesian spatio-temporal models and evaluating their goodness of fit using the Watanabe Akaike Information Criterion (WAIC). Similar research has also been conducted by Ibeji, Mwambi [8], Wahyuni and Syam [9], and Yoo and Wikle [10], showcasing the growing interest in Bayesian spatio-temporal modeling in different contexts.

Previous studies [1, 3] have employed spatio-temporal modeling from a frequentist perspective, driven by limitations such as the absence of subjectivity, challenges in handling prior information, and issues with non-robustness. Additionally, the difficulties in establishing efficient and reliable computing techniques for Intrinsic Conditional Autoregressive (ICAR) models, as outlined by Ver Hoef, Peterson [11], might have prompted the adoption of a spatio-temporal perspective, as seen in works by Fenta, Zewotir [12] and Sahu [13]. Are all spatio-temporal models robust to spatial and temporal correlations [14]? Hence, our goal was to conduct a comparative investigation into the impact of spatio-temporal correlation on four models [13] by evaluating their goodness of fit and biases in modeling food security in Africa. The Bayesian framework allows us to incorporate subjective insights, dealing adeptly with prior information and ensuring robustness—a notable improvement over the limitations identified in frequentist methodologies [11]

To assess goodness of fit, we employed the Watanabe-Akaike Information Criterion (WAIC) for model comparison. This criterion, based on the entire posterior distribution, strikes a balance between model fit and complexity, addressing concerns of both underfitting and overfitting. WAIC is specifically designed to estimate out-of-sample predictive accuracy, making it a versatile and valuable tool for researchers working with various types of Bayesian models [14]. Furthermore, we examined the extent of bias within this category of models through the utilization of Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) measurements using Monte Carlo Simulations.

This comparison helps to understand how each model performs and whether certain models are more robust to variations in the spatio-temporal correlations. By achieving these objectives, the study sought to provide valuable insights into the suitability and performance of various spatio-temporal models in the context of food security analysis, especially considering the impact of different spatio-temporal correlations. The findings would contribute to a better understanding of the strengths and limitations of different modeling approaches and aid in making informed decisions when analyzing food security data in Africa.

Our methodology uniquely combines dimensionality reduction through PCA, Bayesian modeling, and simulation-based exploration, addressing the complexities of spatio-temporal correlation structures in the context of food security. This enables the exploration of diverse policy scenarios, ranging from weak to strong spatio-temporal correlation structures. By simulating these scenarios, we not only account for uncertainties but also offer a comprehensive understanding of potential impacts on the spatio-temporal dynamics of food security in Africa.

The simulation-based approach not only enhances our comprehension of intricate spatio-temporal correlation structures but also equips policymakers with a potent tool for scenario analysis. Policymakers can now evaluate the potential impact of various interventions, spanning from weak to strong spatio-temporal correlation structures. This insight proves invaluable for devising evidence-based policies that effectively address the dynamic challenges of food security in Africa. Given the intrinsic connection between food security, nutritional well-being, and community health, the simulation of diverse scenarios offers a nuanced understanding of potential outcomes. This, in turn, assists public health professionals in designing targeted interventions to enhance the nutritional landscape in the region. Our approach establishes a robust framework for modeling complex spatio-temporal correlation structures.

This paper is structured as follows: In the next section, we describe briefly the methodology of the spatio-temporal interaction models. In Sect. "[Simulation study](#)", we describe the Monte Carlo simulation scheme that was used to examine the performance of the four considered models. In particular, we report on goodness of fit (WAIC) and BIAS (RMSE and MAE). The simulation results are presented in Sect. "[Results](#)". Finally, in Sect. "[Discussion](#)", we summarize our findings and give final remarks.

Methodology and simulation

DATA

The Food and Agriculture Organization (FAO) of the United Nations plays a crucial role in supplying data and information to support the attainment of Sustainable

Development Goal 2. This goal aims to eradicate hunger, food insecurity, and malnutrition on a global scale. The comprehensive dataset on food and nutrition security can be accessed on the FAO website, where you can also find metadata specifically pertaining to Africa. This metadata comprises variable definitions, data sources, years, and units. To address missing values within the dataset, we employed the missForest, which is random forest algorithm [15]. This research covered a 20-year timeframe, ranging from 2000 to 2019. The study gathered a total of 1080 observations, representing data from all 54 countries on the continent. This dataset encompassed 42 variables, each dedicated to specific aspects of food security and nutrition in accordance with the FAO's definitions. The FAO employs the Food Insecurity Experience Scale (FIES) as a metric for measuring food insecurity. In order to maintain consistency and facilitate comparisons between countries, the FIES Survey Module is implemented among nationally representative samples of the adult population. This standardized approach ensures that data on food insecurity is collected in a uniform manner across different nations [16].

Afridi, Jabbar [16] highlighted the significance of utilizing a convergence of evidence strategy, which involves the use of multiple metrics, to identify the key factors associated with food security and nutrition. The initial set of 40 variables sourced from the FAO dataset underwent Principal Components Analysis (PCA). This method was employed to mitigate potential data loss and tackle the challenge of multicollinearity. It also aimed to overcome limitations associated with the conventional selection of based solely on experts' judgment regarding covariates or factors influencing food security and nutrition. PCA offers a more data-driven and objective approach to identifying relevant covariates (explanatory variables) for food security and nutrition. In the end, a total of ten factors were chosen as explanatory variables in the study. These factors included nutrient intake, average food supply, consumption status, childcare, caloric losses, environment, undernourishment, food or nutritional stability, adequate dietary supply, and newborn feeding practices. Together, these ten factors accounted for approximately 74.6% of the total variance present in the dataset [17].

The results of the Kaiser–Meyer–Olkin and Bartlett's Sphericity test values (with a KMO value of 0.729 and a p value of 0.00) indicated that PCA was a suitable approach, as suggested by [18]. Principal component extraction considered eigenvalues larger than one [19]. To ensure that each major component offered distinct information, we applied a VariMax-orthogonal rotation method, in line with Kaiser [19].

The first principal component (PC1) essentially represents nutrient intake and accounts for 19.26% of the total variation in the data. The second principal component (PC2) is associated with factors like food supply, food production, and dietary energy and protein supply. PC2 explains 14.96% of the variation and quantifies the average food supply in Africa. PC3 reflects African consumption levels based on indicators like Gross Domestic Product (GDP) per capita, childhood stunting and overweight, and adult obesity prevalence. It accounts for 10.96% of the variation. PC4, explaining 5.71% of the variance, pertains to childcare factors. PC5 (5.51%) relates to calorie loss, PC6 (4.47%) to environmental influences, PC7 (3.72%) to undernourishment, PC8 (3.43%) to food stability, PC9 (3.39%) to dietary adequacy, and PC10 (3.19%) to infant feeding practices.

The modeling formulation employed in this study is from the spatiotemporal framework. This framework integrates a spatial conditional autoregressive (CAR) prior and

an autoregressive (AR) process. By incorporating both spatial and temporal interconnections, this method effectively captures the underlying patterns in the data. The CAR prior is employed to represent the spatial component, taking into account the spatial autocorrelation between neighboring locations. On the other hand, the AR process is utilized to model the temporal component, capturing the temporal autocorrelation over time. By combining these two components, the model can effectively analyze the spatio-temporal dynamics of the data [20].

In Bayesian statistics, when specifying a prior distribution for random variables exhibiting spatial autocorrelation, a common strategy involves combining a uniform prior that encompasses a wide range of values for the intercept (mean) with the intrinsic conditional autoregressive (ICAR) distribution. This approach allows for flexibility and accommodates the spatial autocorrelation structure in the data. Leroux, Lei [21] proposed a remarkable modification to conditional autoregressive (CAR) models that we incorporated into our work. Their suggestion involves using the equation $Q(\mathbf{W}, \rho) = \rho(\mathbf{W}_d - \mathbf{W}) + (1 - \rho)\mathbf{I}$, Where, \mathbf{W}_d represents a diagonal matrix of the weighted average value derived adjacency, and \mathbf{W} is the n -dimensional adjacency matrix. \mathbf{I} represent the identity matrix of size $N \times N$, and $\mathbf{1}$ is an $N \times 1$ vector of ones. The parameter ρ , which controls the spatial correlation strength, takes on values between 0 and 1, inclusive. This results in a CAR distribution $\varphi_i | \rho, k^2, \varphi_j, j \neq i \sim N\left(\frac{\rho \sum_{j=1}^n w_{ij} \varphi_j}{\rho \sum_{j=1}^n w_{ij} + 1 - \rho}, \frac{k^2}{\rho \sum_{j=1}^n w_{ij} + 1 - \rho}\right)$. In this context, w_{ij} refers to the element in the n th dimensional adjacency matrix \mathbf{W} corresponding to the i th row and j th column. The random effects have a spatial variance of k^2 and autocorrelation of ρ

The spatial weight matrix related to the units i and j is represented by each item $w_{ij} \in \mathbf{W}$. The constituent of w_{ij} is $((i, j))$, which remains the neighborhood matrix with 54×54 dimension. The matrix's nonzero entries reveal whether the two places are neighbors. Typically, the weighted matrix is written as:

$$s_{ij} = \begin{cases} 1 & \text{if areas } i \text{ and } j \text{ are neighbours} \\ 0 & \text{otherwise} \end{cases}$$

The dataset's spatial autocorrelation is verified by applying Moran's $I = \frac{n}{s_0} \frac{\sum_{ij} (w_{ij}(x_i - \mu)(x_j - \mu))}{\sum_i (x_i - \mu)^2}$, 'n' represents the count of points under investigation, x_i, x_j denote the observed values at two distinct points, μ denote the expected value of 'x' and w_{ij} signify the spatial weight element. Moran's I range $[-1, 1]$, large values of the relevant metrics are close to other large value clusters when the value is 1, while large values are close to low values when the value of the Moran's I is -1 [12, 13]

In the model, Y_{it} represented the number of severe food insecurity individuals in a country i ($i = 1, 2, \dots, 54$) during year t ($t = 1, 2, \dots, 20$) and n_{it} denoted the population size of country i at time t . The design matrix \mathbf{X} represented the food security and nutrition covariates components) are represented by the design matrix \mathbf{X} , and β was a vector comprising related fixed effects parameters:

$$\log(\mu_{it}) = \eta_{it} = \mathbf{X}\beta + \log(n_{it}) \tag{1}$$

By introducing the spatio-temporal random effect ψ_{it} to Eq. (1), we obtained the following expression:

$$\log(\mu_{it}) = \eta_{it} + v_{it} \tag{2}$$

To capture the various spatio-temporal features (spatio-temporal interactions) of the multiple aspects of the spatio-temporal random effect (v_{it}), we decomposed it, leading to the four models in the sequel.

Spatio-temporal poisson linear trend model (SPLTM)

$$\log(\mu_{it}) = \eta_{it} + \omega_1 + a_i + (\omega_2 + b_i) \frac{t - \bar{t}}{T} \tag{3}$$

where \bar{t} is calculated as $\frac{T+1}{2}$. Here ω_1 and ω_2 represent the overall intercept and slope parameters, respectively, and are assigned a flat prior distribution. On the other hand, the incremental trend(slope) and intercept parameters for the i th region, denoted as a_i and b_i

respectively, are assigned a conditional autoregressive (CAR) prior distribution with different values of ρ and k^2 . Specifically, the parameters $\mathbf{a} = (a_1, \dots, a_n)$ and $\mathbf{b} = (b_1, \dots, b_n)$ follow the CAR prior distributions $\mathbf{a} \sim \text{CAR}(\mathbf{a} | \rho_{int}, k^2_{int}, \mathbf{W})$ and $\mathbf{b} \sim \text{CAR}(\mathbf{b} | \rho_{slo}, k^2_{slo}, \mathbf{W})$ respectively. Similarly, k^2_{slo} , and k^2_{int} represent the variance parameters. The parameters ρ_{slo} , and ρ_{int} are assigned independent uniform prior distributions in the unit interval (0, 1), indicating that their values can range between 0 and 1. On the other hand, the variance parameters, k^2_{slo} , and k^2_{int} follow the inverse gamma prior distribution. In this model, important trends for severe food insecurity are accounted for and their level of significance is accessed.

Spatio-temporal poisson anova model (SPAM).

Knorr-Held [22] proposed a method for analyzing data that incorporates the interaction between space and time by using a model that factors spatial and temporal main effects based on the Analysis of Variance. Equation 4 is derived from Eq. 2 to represent the spatial and temporal effect with attraction.

$$v_{it} = a_i + b_t + c_{it}, i = 1, \dots, n, t = 1, \dots, T \tag{4}$$

The three sets of parameters indicated by a_i , b_t and c_{it} in Eq. 4 are all considered as random effects, each having its own unique probability distribution.

$$\mathbf{a} | \rho_s, k^2_s, \mathbf{W} \sim \text{CAR}(\mathbf{a} | \rho_s, k^2_s, \mathbf{W}),$$

$$\mathbf{b} | \rho_T, k^2_T, \mathbf{Z} \sim \text{CAR}(\mathbf{b} | \rho_T, k^2_T, \mathbf{Z}),$$

$$c_{it} \sim N(0, k^2_I), i = 1, \dots, n, t = 1, \dots, T,$$

Here \mathbf{Z} is the $T \times T$ matrix that represents temporal adjacency, where each element z_{ij} takes a value of 1 if the absolute difference between i and j is equal to 1, and 0 otherwise. It is assumed that the interaction effect (c_{it}) is independent for all values of i and t . The parameters ρ_s and ρ_T are assigned uniform priors with a range of values between 0 and

1, similar to before(Sect. "DATA"), while the variance parameters k^2_s, k^2_T , and k^2_I are assigned inverse gamma priors.

Spatio-temporal poisson separable model(STSM)

The separable model, which includes an overall time trend with temporal-specific spatial effects, is an alternative model to the one presented in Sect. "Spatio-temporal poisson anova model (SPAM)". In this case, independent conditional autoregressive (CAR) models are assigned to the spatial effects $\mathbf{a} = (a_{1t}, \dots, a_{nt})$ for each $t = 1, \dots, T$ and to $\mathbf{b} = (b_1, \dots, b_T)$, as shown in the following equation.

$$q_{it} = a_{it} + b_t, i = 1, \dots, n, t = 1, \dots, T \tag{5}$$

$$\mathbf{a}_t | \rho_s, k^2_t, \mathbf{W} \sim CAR(\mathbf{a} | \rho_s, k^2_t, \mathbf{W}),$$

$$\mathbf{b} | \rho_T, k^2, \mathbf{Z} \sim CAR(\mathbf{b} | \rho_T, k^2, \mathbf{Z}),$$

\mathbf{Z} refers to the same definition as previously provided in Sect. "Spatio-temporal poisson anova model (SPAM)". The variance parameters k^2_t , where t ranges from 1 to T, and k^2 , are assigned inverse gamma prior distributions, similar to the previous model. The parameters ρ_s and ρ_T are assigned independent uniform prior distributions with a range of values between 0 and 1.

Poisson temporal model for spatio-temporal effect (TMS)

The postulation is made that a particular scenario within the framework of the separable model, as discussed in Sect. "Spatio-temporal poisson separable model(STSM)", corresponds to a temporal autoregressive model with a lag of one. In this special case, the parameter b_t is equal to zero for all values of t, resulting in the simplification as $v_{it} = a_{it}$, here $b_t = 0$ aimed at the entire t and

$$\mathbf{a}_t | \mathbf{a}_{t-1}, \mathbf{W} \sim N(\rho_T, \mathbf{a}_{t-1}, k^2 \mathbf{Q}(\mathbf{W}, \rho_s)^{-1}), t = 2, \dots, T$$

$$\mathbf{a}_1 | \mathbf{W} \sim N(0, k^2 \mathbf{Q}(\mathbf{W}, \rho_s)^{-1}),$$

The precision matrix $\mathbf{Q}(\mathbf{W}, \rho \mathbf{S})$ with spatial dependence is defined in Eq. (1), where the temporal autocorrelation is caused by the mean ρ_T, \mathbf{a}_{t-1} . The prior distributions are assumed to remain unchanged.

The chain models

After decomposing the spatio-temporal random effect v_{it} and adding it to the linear predictor η_{it} as shown in Eqs. 2, four different models were obtained: STLTM, SPAM, STSM, and TMS, represented by Eqs. 6–9, respectively:

$$\log(\mu_{it}) = \eta_{it} + \omega_1 + a_i + (\omega_2 + b_i) \frac{t - \bar{t}}{T} \tag{6}$$

$$\log(\mu_{it}) = \eta_{it} + a_i + b_t + c_{it} \tag{7}$$

$$\log(\mu_{it}) = \eta_{it} + a_{it} + b_t \tag{8}$$

$$\log(\mu_{it}) = \eta_{it} + a_{it} \tag{9}$$

Markov chain monte carlo: hamiltonian monte carlo

To construct a Markov chain that converges to the target distribution, Markov Chain Monte Carlo (MCMC) techniques are employed. A Markov chain is constructed by iteratively sampling $\mathbf{q}^{(h+1)}$ from the conditional distribution $p(\mathbf{q}|\mathbf{q}^h)$ at each time point $h > 0$. The process begins with an initial value $\mathbf{q}^{(0)}$. Importantly, the value $\mathbf{q}^{(h+1)}$ at each step depends solely on the current value $\mathbf{q}^{(h)}$, and is independent of its more distant past values $\mathbf{q}^{(h-1)}, \mathbf{q}^{(h-2)}, \dots, \dots, \mathbf{q}^{(0)}$. This specific property, where each step in the chain only depends on the current state and not on previous states beyond the immediate predecessor, is a fundamental requirement for the chain $\mathbf{q}^{(h)}$ to be considered a first-order Markov Chain.

The primary challenge in Bayesian modeling convergence lies in constructing a Markov chain that possesses essential properties such as stationarity, irreducibility, aperiodicity, and ergodicity, while also having a stationary distribution equal to the target posterior distribution $\pi(\mathbf{q}|\mathbf{y})$. Although there are alternative methods like the Metropolis–Hastings algorithm and the Gibbs sampler, the Hamiltonian Monte Carlo (HMC) is widely regarded as the superior choice. This is because HMC excels in exploring regions of high posterior probability and concurrently employs a Metropolis–Hastings step to address the issue of oversampling peaks while neglecting low probability areas. For a comprehensive understanding of HMC, numerous excellent resources are available, including Lambert [23]. So, executing Hamiltonian Monte Carlo (HMC) for a parameter vector θ requires following the procedures delineated in Lambert’s 2018 publication, which are presented in a sequel.

1. Choose an initial location, $\mathbf{q}^{(0)}$, randomly from an initial proposal distribution.
2. During iteration h , generate an initial momentum (\mathbf{m}) randomly from a proposal distribution, such as $\mathbf{m} \sim N(\boldsymbol{\mu}, \boldsymbol{\Sigma})$.
3. Progress the current point \mathbf{q}, \mathbf{m} to $\mathbf{q}^*, \mathbf{m}^*$ by executing L steps of the leapfrog algorithm, as implemented by the No-U-Turn Sampler (NUTS) [24]. The pair $\mathbf{q}^*, \mathbf{m}^*$ represents the updated position of the parameter and its momentum after completing the L leapfrog steps.

4. Compute the Metropolis acceptance probability, denoted as $\alpha(\mathbf{q}, \mathbf{g})$, for the proposal $\mathbf{g} = (\mathbf{q}^*, \mathbf{m}^*)$ and current point $\mathbf{q} = (\mathbf{q}, \mathbf{m})$ based on the target density $\pi(\mathbf{q}, \mathbf{m}|\mathbf{y})$ where

$$\log\pi(\mathbf{q}, \mathbf{m}|\mathbf{y}) \propto \log(\pi(\mathbf{q}|\mathbf{y})) + \frac{1}{2}\mathbf{m}'\Sigma^{-1}\mathbf{m}.$$

5. Generate a random number u from a uniform distribution $U(0,1)$, and set $\mathbf{q}^{(k+1)} = \mathbf{q}^*$ if $u < \alpha(\mathbf{q}, \mathbf{g})$ otherwise, set $\mathbf{q}^{(k+1)} = \mathbf{q}$

We employed the MCMC algorithm that generates 10,000 samples from the posterior distribution with $N = 120,000$, $\text{burn.in} = 20,000$, and $\text{thin} = 10$.

Simulation study

The main aim of our study was to explore how the spatial and temporal correlation structures impact the spatio-temporal dynamics of food security and nutrition in Africa. To achieve this, we conducted a Monte Carlo simulation involving 1,000 random samples (replications) with varying spatial and temporal correlations (0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95) for longitudinal twenty repeated measures in each country. Every aspect of the analysis, which encompassed tasks such as data generation, simulation, and the estimation of model parameters, was executed utilizing the R programming language (R 4.3.1). For each scenario or combination of parameter values, we executed a total of 1000 simulation runs.

The combination of PCA and VIF allows for a more holistic evaluation of multicollinearity, validating the efficacy of the PCA dimensionality reduction while ensuring the dependability of regression model outcomes. This integrated approach fortifies the robustness of statistical analyses, contributing to a more comprehensive understanding of the interrelationships among variables, as evidenced in prior studies [25, 26]. From Fig. 1, we observed that all ten components had a VIF value of 1. The numerical value

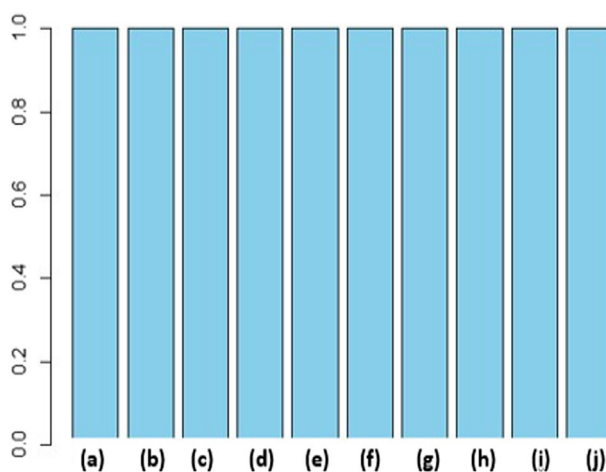


Fig. 1 Variable Inflation Factor (VIF) for the 10 explanatory factors of FSN in Africa: **a** nutrient intake; **b** average food supply; **c** consumption status; **d** childcare; **e** caloric losses; **f** environment; **g** undernourishment; **h** food or nutritional stability; **i** adequate dietary supply; **j** newborn feeding practices

of 1 for VIF observed for these 10 components confirms that there is no inflation in the percentage of variance (i.e., the standard error squared) for each coefficient. Therefore, the application of PCA is justified.

Data-generating process

During the data generation process (response variable), we specified four versions of spatio-temporal models to investigate the effects of spatio-temporal correlation using randomly generated (artificial) dependent variables and the 10 explanatory variables (covariates) of food security in Africa identified when the 40 indicators from FAO was subjected to PCA. The spatial units were the 54 countries of Africa. These ten components (covariates) were obtained from the 40 variables provided by the FAO dataset after subjecting them to Principal Component Analysis (PCA) as described in Sect. "[Methodology and simulation](#)".

We devised the following iterative algorithm, which was utilized to induce the desired spatio-temporal correlation when generating the values of the dependent variables.

The first version.

- a. Define the spatial correlation (0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95) and temporal correlation value (0.05, 0.15, 0.25, 0.35, 0.45, 0.55, 0.65, 0.75, 0.85, 0.95) combinations, one hundred pairs.
- b. Use information in (a) to generate two correlation matrices: the 'spatial correlation matrix' and 'temporal correlation matrix'. These matrices have the specified correlation values and dimensions based on 'num_countries' and 'num_observations', respectively.
- c. Use the information in (b) to generate the noise using multivariate normal distributions with means of 0 and covariance matrices 'spatial_correlation_matrix' and 'temporal_correlation_matrix', in(b). Here the spatial and temporal data is generated by adding the spatial noise to the coordinates and replicating the temporal noise.
- d. Use information (c) to calculate a latent variable $\lambda = \exp(\text{spatial data} + \text{temporal data})$. Where 'spatial data' represents the x-coordinate of the spatial data for each country. 'temporal data': This represents the temporal noise generated using multivariate normal distributions.
- e. Generate the observed response variable y where each count is expected to be around the corresponding ' λ ' value. Where $y > 0$
- f. Fit spatio-temporal interaction model for the four (SPLTM, SPAM, STSM, TMS) models specified in Sect. "[The Chain Models](#)"
- g. Estimate the parameters, and compute the bias and goodness-of-fit measures for evaluation
- h. Perform steps (a) to (g) iteratively for a total of 1,000 repetitions, each time varying the spatial correlation and temporal correlation values. After each iteration, calculate the relevant parameters accordingly.

For model validation:

- I. All other procedures from (a) to (e) remained unchanged, as outlined in the first version.
- II. To enable the computation of validation statistics, designate 70 percent of the data for training and 30 percent for testing.
- III. Fit spatio-temporal interaction model for the SPAM model
- IV. Estimate the parameters, and compute the coverage percentage for validation
- V. Perform steps (I) to (IV) iteratively for a total of 1,000 repetitions, each time varying the spatial correlation and temporal correlation values. After each iteration, calculate the relevant parameters accordingly.

Statistical analyses in simulated datasets

To determine the best predictive model from the four available options, we employed a range of statistical metrics to evaluate and compare their performance. Specifically, we computed the WAIC, RMSE, and MAE for each of these models. The model with the lowest statistical metric values indicates superior predictive ability, striking a balance between model fit and complexity.

Goodness of fit test: WAIC

The optimal fit among the four competing models was evaluated using WAIC [27, 28]. To calculate WAIC, the log-likelihood of each data point is computed based on the model's parameters. The total log-likelihood is then adjusted for the variance of these log-likelihoods, with a correction term for effective parameters.

$$\text{WAIC} = -2(\text{lpd} - p\text{WAIC}) \quad (10)$$

Here, "lpd" represents the log pointwise predictive density, which is calculated as the summation from i equals 1 to n of the natural logarithm of the density function $p(y_i|\theta)$. In this context, $p(y_i|\theta)$ represents the density of observing the data point y_i given the model parameters θ . The effective number of parameters, denoted as "pWAIC," is computed as the summation from i equals 1 to n of the variance of the log predictive densities, $\text{Var}[\log\{p(y_i|\theta)\}]$. We assessed the goodness of fit by calculating the average model-based WAIC across the 1000 replications.

BIAS: RMSE and MAE

Let \hat{y}_i denote the estimated severe food security value for the i -th observation using a specific method, while y_i represents the observed value for the same i -th observation. The variable n represents the number of data points in each replication. The RMSE is calculated by taking the square root of the average of the squared differences between the actual values and the estimated parameters generated by the model in each iteration. The RMSE can be expressed as follows:

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (11)$$

The MAE is determined by taking the mean of the absolute differences between the actual values and the estimated parameters obtained from the model in each iteration, as described by Eq. 12.

$$\text{MAE} = \left[\frac{\sum_{i=1}^n |y_i - \hat{y}_i|}{n} \right] \quad (12)$$

Both RMSE and MAE for each replication are arithmetically averaged throughout the 1000 replication.

Validation

This involved the validation of the best model among the four competing models. We computed the coverage percentage (CP) for the 95% predictive intervals using the formula $CP = 100 \frac{1}{k} \sum_{i=1}^k I(L_i \leq y_i \leq U_i)$, where y_i represents the observed value for $i = 1, \dots, k$ and (L_i, U_i) represents the $100(1 - \alpha)\%$ predictive interval for predicting y_i . In this calculation, we employed the indicator function $I(\cdot)$. It's important to note that we used 70% of the data as the training set and the remaining 30% as the test set [29, 30].

Results

We present the simulation results based on the types of estimated statistical metrics explained in the earlier section. Our primary focus is on evaluating the performance of spatio-temporal models for food security concerning Africa. These metrics are visually represented in a matrix plot and Table.

For the goodness of fit test, it is represented in Table 1. In Table 1 “SP” represents spatial correlation, “TM” denotes temporal correlation, or the correlation over time. The cells in the table are colored-coded to easily identify the best and worst results. Red cells indicate higher or worse values, meaning WAIC. Yellow cells point to lower or better values, signifying less WAIC. The color scheme facilitates interpretation by signaling which areas or spatio-temporal correlation pairs exhibited the strongest or weakest WAIC metric.

Matrix plots are a crucial tool for presenting the relationships and patterns discovered in our analysis of the four competing spatio-temporal models. They provided a visual narrative that illustrates how different spatial and temporal correlation structures impact food security and nutrition modeling for Africa. In these matrix plots, the cells are color-coded to represent the values of the parameters, including RMSE, and MAE, for each of the four models under consideration. lower RMSE and MAE values (deeper colors) indicate better model performance. This visualization method helps us to grasp and compare the biases of these models across various spatial and temporal correlation scenarios.

Table 1 (continued)

SPLTM	10572.15	10194.71	11884.68	12402.68	12183.83	14597.93	11663.72	11468.87	15179.53	10404.57
TMS	4497.50	4357.75	4268.05	4347.83	4353.34	4326.43	4410.11	4359.09	4540.93	4374.86
SPAM	4203.14	4260.74	4646.11	4481.25	4525.17	4409.73	4418.61	4244.16	4479.40	4339.27
STSM	4657.03	4396.43	4563.45	4593.52	4694.66	4816.72	4463.89	4702.84	4849.61	4914.86
	SP = 0.65	TM = 0.05	SP = 0.65	TM = 0.15	SP = 0.65	TM = 0.25	SP = 0.65	TM = 0.35	SP = 0.65	TM = 0.45
SPLTM	9486.1	8421.3	12149.2	12269.4	12713.0	16823.6	11646.7	13204.9	15455.3	17685.0
TMS	4475.35	4335.66	4442.67	4505.57	4409.91	4277.66	4473.67	4569.30	4690.93	4615.30
SPAM	4422.23	4376.55	4174.75	4295.12	4383.27	4297.17	4389.86	4417.35	4467.96	4410.75
STSM	4609.64	4429.22	4539.38	4552.63	4813.25	4799.23	4437.24	4536.05	4674.10	4838.36
	SP = 0.75	TM = 0.05	SP = 0.75	TM = 0.15	SP = 0.75	TM = 0.25	SP = 0.75	TM = 0.35	SP = 0.75	TM = 0.45
SPLTM	12174.4	10129.5	8315.4	12475.9	1529.1	13067.6	13273.4	12595.5	16834.5	14705.1
TMS	4285.76	4460.81	4430.19	4437.87	4244.78	4365.72	4658.00	4257.56	4672.68	4461.23
SPAM	4451.70	4289.98	4400.27	4378.13	4533.67	4311.52	4594.10	4252.16	4276.54	4255.69
STSM	4460.54	4406.18	4455.06	4808.25	4703.10	4709.35	4622.54	4607.99	4574.17	4519.13
	SP = 0.85	TM = 0.05	SP = 0.85	TM = 0.15	SP = 0.85	TM = 0.25	SP = 0.85	TM = 0.35	SP = 0.85	TM = 0.45
SPLTM	9591.1	8590.2	10832.5	9896.1	7158.5	12564.8	11052.4	10916.6	11440.9	11652.9
TMS	4345.51	4563.08	4467.29	4502.78	4497.35	4338.49	4479.49	4404.09	4659.02	4495.42
SPAM	4505.85	4343.92	4197.17	4383.23	4474.18	4426.53	4497.29	4711.44	4418.01	4550.60
STSM	4443.47	4446.04	4646.95	4464.84	4713.39	4601.12	4918.00	4624.68	4557.76	4511.20
	SP = 0.95	TM = 0.05	SP = 0.95	TM = 0.15	SP = 0.95	TM = 0.25	SP = 0.95	TM = 0.35	SP = 0.95	TM = 0.45
SPLTM	11449.4	10088.7	13495.1	17701.4	14291.9	14274.3	16134.1	11912.1	18028.3	14910.5
TMS	4281.91	4450.83	4298.99	4342.94	4060.26	4435.40	4238.69	4371.53	4479.40	4548.91
SPAM	4458.82	4359.0	4459.8	4096.4	4420.3	4481.1	4190.5	3893.4	4469.2	4503.28
STSM	4596.4	4377.44	4407.07	4212.43	4527.54	4544.56	4419.91	4636.35	4521.10	4737.9

In the table, "SP" represents spatial correlation, and "TM" represents temporal correlation. Cells highlighted in red indicate worse performance, while yellow cells indicate better performance

Model (SPLTM) exhibits values ranging from 6000 to 20,000, with the model exhibiting the highest cluster of WAIC values when compared to the four competing models. The highest WAIC value observed was 18,450.05, while the lowest value recorded was 6529.104. Notably, lower WAIC values were observed when both spatial and temporal correlations were weak to moderate, falling within the range of 0.05 to 0.5. This suggests that the model performs better or exhibits more favorable characteristics under these conditions. Conversely, higher WAIC values were observed in scenarios characterized by strong temporal and spatial correlations. This implies that the model's performance does not improve with higher values of the parameters, specifically in cases of elevated spatial and temporal correlation. This behavior indicates that the model is particularly

sensitive to increased spatial and temporal correlation, particularly when there are high levels of spatio-temporal correlation.

For spatio-temporal Poisson Anova Model (SPAM), the WAIC values are distributed within the range of 3800 to 5000. The lowest WAIC value observed was 3893.43, while the highest recorded WAIC value was 4917.99 (Table 1). Among the four models compared, SPAM consistently obtained the lowest WAIC scores in the vast majority of simulated scenarios concerning diverse spatial and temporal dependency combinations. This indicates that SPAM excelled in capturing and modeling the underlying spatial and temporal correlations within the data, as determined by the WAIC metric for model selection and performance. In the case of the Poisson Temporal Model (TMS) for the Spatio-temporal correlation effect, as shown in Table 1, the WAIC values also fall within the range of 4000 to 5000 for all 100 pairs of spatio-temporal correlation. The lowest WAIC value was 4060.257, and the highest WAIC value was 4756.981.

It is worth noting an interesting pattern observed in the results: lower WAIC values, indicating better model fit, were seen when the spatial and temporal correlation values included in the simulations ranged from low to moderately strong, specifically when both types of correlations took on values between 0.05 and 0.75. This pattern held for both the SPAM and TMS models. In other words, these two models performed best at capturing the relationships in the data when the spatial and temporal dependencies were in the moderate rather than extremely low or high ranges. This provides useful insight into the conditions under which these types of spatiotemporal models are better able to model correlation structures. WAIC values deteriorated when the data exhibited strong spatio-temporal correlations, as opposed to when these correlations were at lower or moderate levels. This implies that although the SPAM and TMS models exhibited good performance when dealing with modest spatial and temporal dependencies, their capability to effectively capture patterns in the data decreased as the strength of the dependencies across geographical areas and over time increased substantially.

Concerning the performance of the spatio-temporal Poisson Separable Model (STSM), a notable observation is that the performance of the STSM is particularly responsive to variations in both spatial and temporal correlation. The STSM demonstrates relatively strong performance in scenarios characterized by low to moderate levels of spatial and temporal correlation, ranging from 0.05 to 0.55. However, it exhibits higher WAIC values when there is a combination of high spatial correlation with lower temporal correlation. The WAIC values in this case fall within the range of 4100 to 5000, with the highest WAIC value recorded at 4999.791 and the lowest at 4150.11.

Bias evaluation

Figure 2 depicts the distribution of bias levels, as indicated by RMSE, for all four models. Notably, for both the SPAM and TMS models, the values fall within the range of 0.7 to 1.2, and interestingly, both models exhibited similar distribution patterns. The parameter values vary according to different levels of spatio-temporal correlation, which ranged from 0.7 to 1.00. Notably, there was only one observation exceeding 1.00, occurring when the temporal correlation was 0.25 and the spatial correlation was 0.05. According to Fig. 2, the SPLTM exhibits the highest range of RMSE values, spanning from 1.2 to

16.94. In contrast, the STSM shows a narrower range, with RMSE values ranging from 1.8 to 0.89.

Regarding the MAE results presented in Fig. 3, the SPLTM model exhibited the widest range of values, spanning from 8.9 to 0.7. In contrast, the SPAM model displayed narrower intervals, with values ranging from 0.8 to 0.5. For the TMS model, the MAE intervals were 0.99 to 0.5 (Fig. 3). Interestingly, the majority of MAE metrics for SPAM fell within the range of 0.6 to 0.75, except in situations when spatial correlation was 0.65 and temporal correlation was 0.05 (Fig. 3). Similar patterns were observed for cases of

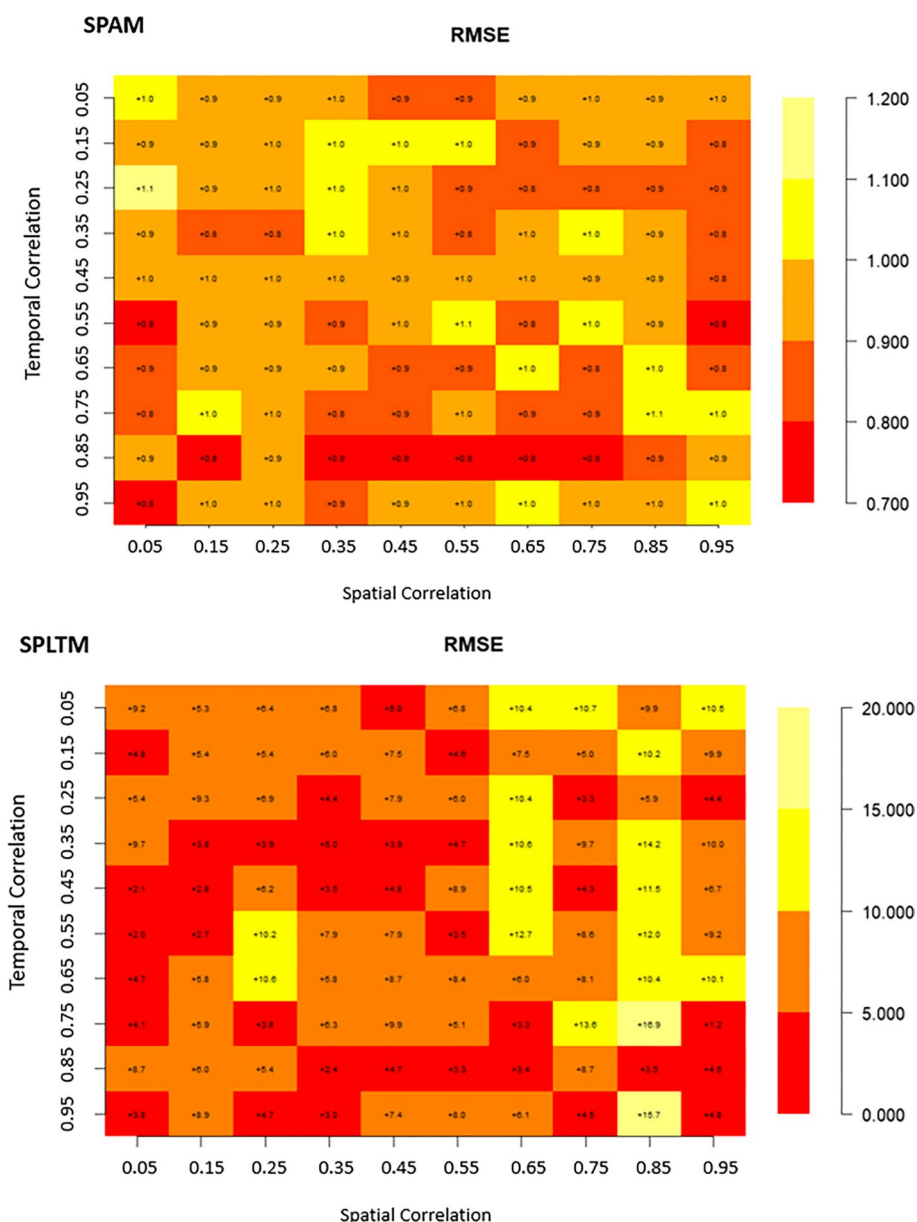


Fig. 2 Bias (RMSE) of spatiotemporal models (SPAM, SPLTM, STSM, and TMS) with varying spatio-temporal correlation

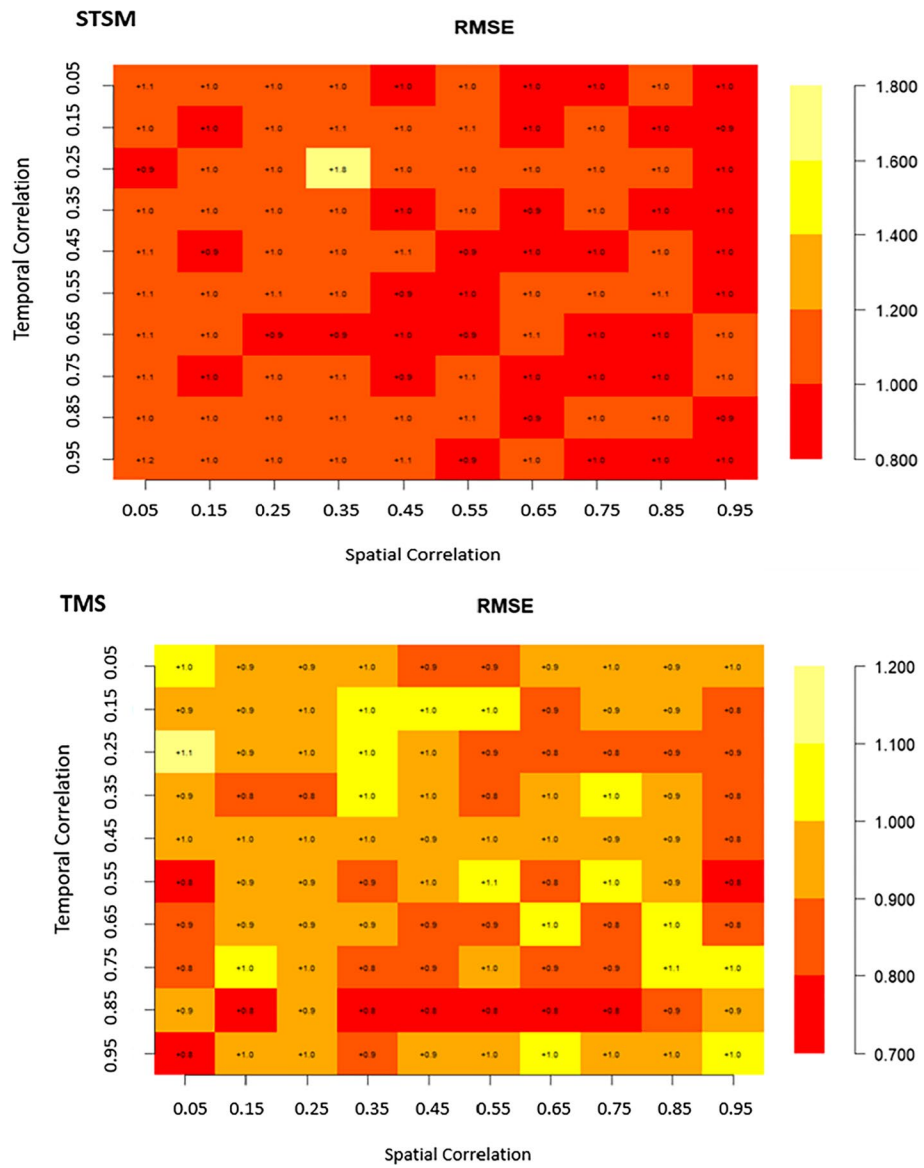


Fig. 2 continued

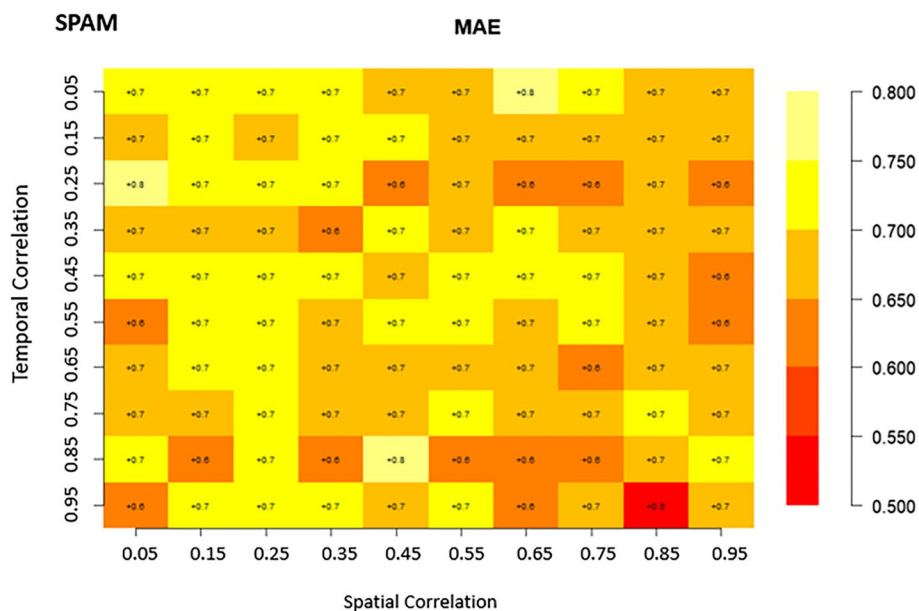


Fig. 2 continued

low spatial (0.05) and temporal (0.25) correlation, as well as spatial (0.45) and temporal (0.85) correlation. These distributions suggest that the SPAM model may not perform well under such conditions, but it excels when there is a strong correlation in both spatial (0.85) and temporal (0.95) (Fig. 3). For the TMS model, lower MAE metrics were clustered when spatial correlation ranged from 0.45 to 0.95 for nearly all temporal correlations (Fig. 3). Regarding the STSM model, the MAE metrics ranged from 0.65 to 0.85, with the highest values observed in scenarios where both spatial and temporal correlations were very weak (0.05) or when temporal correlation was strong (0.95) with weak spatial correlation (0.05) (Fig. 3).

Coverage percentage-based validation of the selected model: SPAM

Based on the principles of statistical learning, the SPAM model is considered the best among the four competing models. Therefore, it is crucial to validate its performance using the 70% train and 30% test split from a machine learning perspective. This validation will help assess the model’s capability for handling unseen data across various spatio-temporal correlation scenarios especially concerning Africa’s food security modelling. We noticed a homogeneous distribution of the coverage percentage across almost all variations in spatio-temporal correlation concerning the SPAM. The coverage percentages were noted to vary between 96.8% and 80.6%. Notably, the SPAM model consistently achieved higher coverage percentages (> 90%) for nearly all levels of spatio-temporal correlation, with one exception: when the temporal correlation was high (0.85) and the spatial correlation was weak (0.05) Fig. 4. It achieved its lowest coverage percentage at 80 (Fig. 4), although there were instances (17 observations) where the coverage percentage ranged from 85 to 90.

Discussion

As we concentrate on evaluating the performance of spatio-temporal models under different spatio-temporal correlation scenarios, we summarize our findings in a table and matrix plots. When necessary, we contextualize our results within the existing literature.

In the table and matrix plot displaying WAIC, RMSE, and MAE values for our four models (SPLTM, SPAM, STSM, and TMS), we have identified intriguing and consistent monotonic patterns that offer valuable insights into their performance and bias characteristics. In essence, these observations underline the sensitivity of all four models to variations in spatio-temporal correlation. On average, each of these models

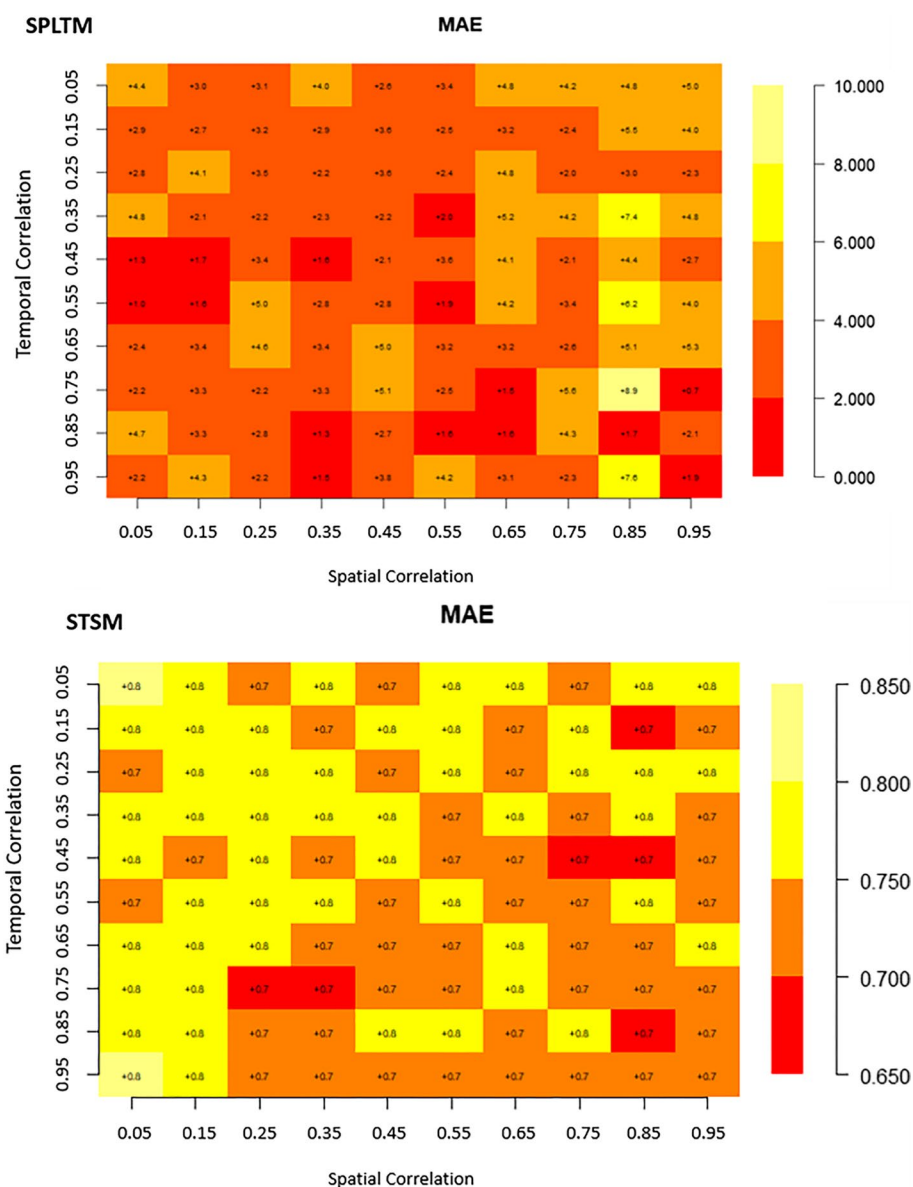


Fig. 3 Bias (MAE) of spatiotemporal models (SPAM, SPLTM, STSM, and TMS) with varying spatio-temporal correlation

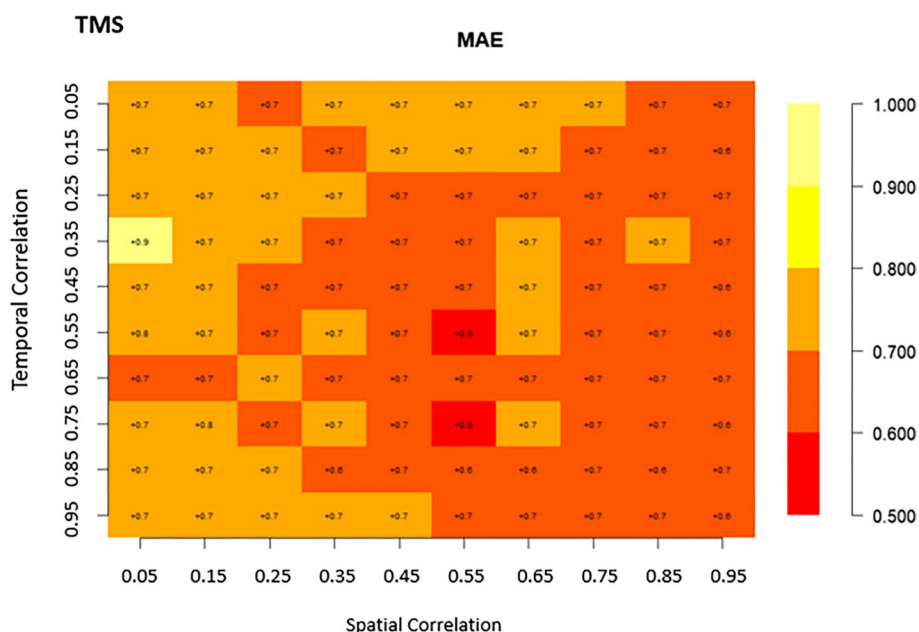


Fig. 3 continued

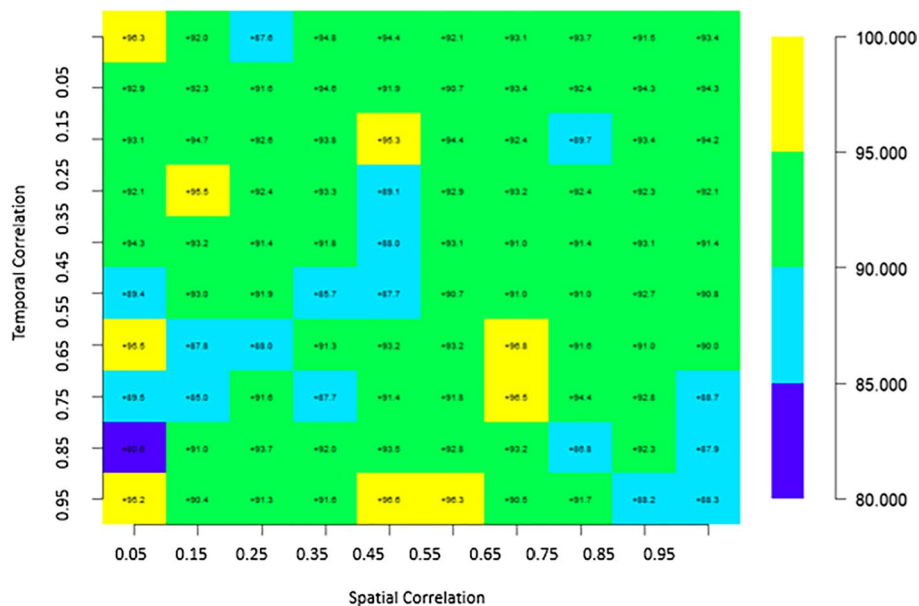


Fig. 4 Distribution of coverage percentage for the SPAM model

gradually converges towards the true values as spatio-temporal correlation increases, which is reflected by the corresponding increase in WAIC values. These findings strongly support the notion that an augmentation in spatio-temporal correlation levels can indeed impact the performance of these models [31]. Our results reinforced the significance of selecting models that align with the specific patterns evident in the data.

Our study highlights that the SPLTM consistently exhibited WAIC values in these comparisons. This recurrent pattern of SPLTM holding the highest WAIC values raises significant concerns about its suitability for modeling food security and nutrition in the African context. In stark contrast, our findings emphasize the remarkable consistency of the SPAM model. SPAM consistently outperformed the other models, as evidenced by its persistently lower WAIC values when compared to SPLTM, STSM, and TMS for almost all scenarios under study. This robust and unwavering performance strongly suggests that SPAM stands out as exceptionally well-suited for the complex task of modeling the dynamic interplay of food security and nutrition in the African context. The reliability demonstrated by SPAM makes it a top choice for researchers seeking to comprehend and predict the intricate spatio-temporal patterns within this critical domain. These findings align with existing research [13] on the behavior of spatio-temporal interaction models and reinforce the superiority of the SPAM model in this context.

Our analysis of bias using RMSE and MAE unearthed interesting patterns that shed light on the behavior of the four spatio-temporal models. The SPLTM model consistently exhibited relatively larger bias values across a wide spectrum of spatio-temporal correlation values. This consistent pattern of overestimation in food security across diverse spatio-temporal scenarios underscores the model's limitations and its tendency to inflate predictions when necessary. In contrast, the Poisson Temporal Model (TMS) demonstrated relatively more variable bias patterns. In scenarios with low spatial and temporal correlation, it consistently recorded substantial bias values, indicating a potential tendency to overestimate food security. However, in high correlation settings, TMS exhibited a notable shift towards underestimation, showcasing the influence of varying spatio-temporal correlation on its predictive behavior.

The SPAM emerged as the model with minimal bias across most scenarios. This consistent performance highlighted SPAM's reliability in accurately modeling food security dynamics, irrespective of the degree of spatio-temporal correlation. Conversely, the STSM exhibited a notably high bias, consistently underestimating food security. This bias pattern was particularly prominent in areas characterized by low to moderate spatial and temporal correlation, emphasizing the model's limitations in accurately capturing the dynamics of food security in such settings. Our bias evaluation not only highlights the relative performance of these models but also provides researchers with critical insights for model selection based on their specific research questions and the characteristics of their spatio-temporal data. Understanding these bias patterns is paramount to making informed decisions when choosing the most appropriate model for accurate and reliable predictions in the domain of food security.

In light of these findings, we recommend that model selection should be a balanced consideration of both biasness and goodness of fit metrics. A model that exhibits superior goodness of fit but introduces significant bias may not provide accurate predictions, and conversely, a model with low bias but poor goodness of fit may fail to capture underlying spatio-temporal dynamics. Researchers should aim to strike a harmonious equilibrium between these two aspects to select models that align with the specific characteristics of their data and research objectives.

The statistical insights gained from our research have the potential to extend their impact beyond the realm of spatio-temporal modeling. Other disciplines grappling with complex and correlated data, such as epidemiology, environmental science, and economics, can draw upon the methodology and lessons learned in this study. Our research fosters interdisciplinary collaboration and opens new avenues for the analysis of intricate spatio-temporal patterns in diverse fields. This cross-disciplinary applicability enhances the relevance and broader impact of our findings.

While the integration of PCA and Bayesian modeling contributes to the efficiency of our methodology, it is crucial to recognize potential computational complexities, particularly with larger datasets. Given that our study focused on 20 temporal observations for each subject (country), future research endeavors could delve into computational optimization techniques. These techniques would be beneficial for handling datasets of varying sizes, ranging from smaller (e.g., 5, 10, 15) to larger (e.g., 25, 30, 35) temporal observations. Implementing such strategies would help maintain computational efficiency without compromising model accuracy, addressing scalability concerns associated with different data sizes. The Bayesian modeling approach relies on specific assumptions, and deviations from these assumptions can influence the robustness of the findings. For example, the assumed prior distributions may introduce subjectivity. Future studies could incorporate sensitivity analyses to evaluate the impact of different priors on model outcomes. This would contribute to enhancing the credibility and reliability of our Bayesian approach.

To the best of our understanding, there hasn't been prior research assessing the performance of the four spatio-temporal models (SPLTM, SPAM, STSM, and TSM) in terms of their sensitivity to spatio-temporal correlations using Monte Carlo simulations. It's important to note that, as with any simulation study, our findings are bound by the specific scenarios and data employed in our simulations. Hence, while our analysis of bias and goodness of fit provides valuable insights into the capabilities and constraints of spatio-temporal modeling approaches, it's important to note that these findings might not be directly transferrable to scenarios or contexts that were not explicitly investigated in our study.

Conclusion

The primary objective of our study was to investigate the influence of spatial and temporal correlation structures on the spatio-temporal patterns of food security and nutrition (FSN) in Africa. We conducted this exploration within the spatio-temporal framework, considering four Bayesian Poisson models: SPLTM, SPAM, STSM, and TSM. Monotonic patterns became evident when we subjected the four competing models to varying spatio-temporal correlations. The SPAM model consistently exhibited the lowest RMSE, MAE, and WAIC values across all spatio-temporal correlations. This notable consistency implies a high level of robustness in the SPAM model's performance. These results strongly suggest that SPAM is a reliable and solid choice for capturing the dynamics of food security. The results of our study provide practical guidance for researchers and practitioners engaged in spatio-temporal modeling. By identifying the strengths and weaknesses of each model under varying degrees of

spatio-temporal correlation, we offer a valuable framework for informed model selection. Researchers should strive to find a balanced equilibrium between these two factors (goodness of fit and bias metrics) when choosing models that are in line with their data's unique features and research goals. This knowledge empowers researchers to select models that offer reliability and consistency, enhancing the applicability of their findings.

Abbreviations

SPLTM	Spatio-temporal poisson linear trend model
TMS	Poisson temporal model
SPAM	Spatio-temporal poisson anova model
STSM	Spatio-temporal poisson separable model
WAIC	Watanabe akaike information criterion
RMSE	Root mean square error
MAE	Mean absolute error
MCMC	Markov chain monte carlo
FAO	Food and agriculture organization
PCA	Principal components analysis
PC	Principal component
GDP	Gross domestic product
CAR	Conditional autoregressive
ICAR	Intrinsic conditional autoregressive
HMC	Hamiltonian monte carlo
SP	Spatial correlation
TM	Temporal correlation

Acknowledgements

We gratefully acknowledge the Food and Agriculture Organization (FAO) for providing valuable data on food security, which greatly contributed to the research presented in this paper. Their dedication to collecting and disseminating reliable information is essential in addressing global food security challenges.

Author contributions

The authors, AB and TZ, were responsible for the conception and design of the experiments, analysis, and interpretation of the data, and writing of the paper.

Funding

This project work was carried out without any external funding (public, commercial, or not-for-profit sectors).

Availability of data and materials

The datasets used and/or analysed during the current study available from the corresponding author on reasonable request.

Declarations

Ethics approval and consent to participate

The Suite of Food Security Indicators presents a fundamental collection of indicators focused on food security. This selection is based on the recommendations of experts who convened during the Committee on World Food Security (CFS) Round Table on hunger measurement at FAO headquarters in September 2011. These indicators were chosen to encompass various dimensions of food insecurity. The selection process considered expert insights and the availability of data that permits cross-regional and temporal comparisons. The FAO Statistical Quality Assurance Framework's Principle 10 underscores the strict confidentiality of data subject to national policies, particularly regarding individuals, entities, and small aggregates. Such data are exclusively used for statistical purposes or in accordance with legal mandates.

Consent for publication

Not applicable for that section.

Competing interests

The authors declare that they have no conflict of interest.

Received: 26 October 2023 Accepted: 18 April 2024

Published online: 27 April 2024

References

1. Nicholson CF, et al. Food security outcomes in agricultural systems models: current status and recommended improvements. *Agric Syst.* 2021;188:103028.

2. Zhai H, et al. Understanding spatio-temporal patterns of land use/land cover change under urbanization in Wuhan, China, 2000–2019. *Remote Sens.* 2021;13(16):3331.
3. Cooper MW, et al. Text mining the food security literature reveals substantial spatial bias and thematic broadening over time. *Glob Food Sec.* 2020;26:100392.
4. Donatelli C, et al. Statistical detection of spatio-temporal patterns in the salinity field within an inter-tidal basin. *Estuaries Coasts.* 2022;45(8):2345–61.
5. Masolele RN, et al. Spatial and temporal deep learning methods for deriving land-use following deforestation: a pan-tropical case study using Landsat time series. *Remote Sens Environ.* 2021;264:112600.
6. Santos-Fernandez, E., et al., *SSNbayes: An R package for Bayesian spatio-temporal modelling on stream networks.* arXiv preprint [arXiv:2202.07166](https://arxiv.org/abs/2202.07166), 2022.
7. Aswi A, Sukarna SC, Mengersen K. Effects of climatic factors on dengue incidence: a comparison of bayesian spatio-temporal models. *J Phys Conf Ser.* 2021;1863:012050. <https://doi.org/10.1088/1742-6596/1863/1/012050>.
8. Ibeji JU, Mwambi H, Iddrisu A-K. Bayesian spatio-temporal modelling and mapping of malaria and anaemia among children between 0 and 59 months in Nigeria. *Malar J.* 2022;21(1):1–12.
9. Wahyuni, M.S. and R. Syam. *Comparison of Bayesian Spatio-temporal Models of Tuberculosis in Makassar, Indonesia.* in *Journal of Physics: Conference Series.* 2021. IOP Publishing.
10. Yoo, M. and C.K. Wikle, *A bayesian spatio-temporal level set dynamic model and application to fire front propagation.* arXiv preprint [arXiv:2210.14978](https://arxiv.org/abs/2210.14978), 2022.
11. Ver Hoef JM, et al. Spatial autoregressive models for statistical inference from ecological data. *Ecol Monogr.* 2018;88(1):36–59.
12. Fenta HM, Zewotir T, Muluneh EK. Space–time dynamics regression models to assess variations of composite index for anthropometric failure across the administrative zones in Ethiopia. *BMC Public Health.* 2022;22(1):1–11.
13. Sahu SK. Bayesian modeling of spatio-temporal data with R. Boca Raton: Chapman and Hall/CRC; 2021.
14. Devine, S., Falk, C.F., Fujimoto, K.A., Comparing the Accuracy of Three Predictive Information Criteria for Bayesian Linear Multilevel Model Selection. 2023
15. Bofa A, Zewotir T. Filling the gap in food and nutrition security data: What imputation method is best for Africa's food and nutrition security? *Lith J Stat.* 2022;61:16–31.
16. Afridi GS, et al. An analysis of food insecurity in pakistan: prevalence of undernourishment (PoU) and food insecurity experience scale (FIES). *J Appl Econ Bus Stud.* 2021;5(1):175–90.
17. Bofa A, Zewotir T. The determinants of severe food insecurity in Africa using the longitudinal generalized Poisson mixed model. *Lith J Stat.* 2023;62:1–19.
18. Bartlett MS. A note on the multiplying factors for various χ^2 approximations. *J R Stat Soc Ser B Stat Method.* 1954;16(2):296–8. <https://doi.org/10.1111/j.2517-6161.1954.tb00174.x>.
19. Kaiser HF. The application of electronic computers to factor analysis. *Educ Psychol Measur.* 1960;20(1):141–51.
20. Jalilian A, Mateu J. A hierarchical spatio-temporal model to analyze relative risk variations of COVID-19: a focus on Spain, Italy and Germany. *Stoch Env Res Risk Assess.* 2021;35:797–812.
21. Leroux BG, Lei X, Breslow N. Estimation of disease rates in small areas: a new mixed model for spatial dependence. In: Elizabeth Halloran M, Berry D, editors. *Statistical models in epidemiology, the environment, and clinical trials.* New York, NY: Springer New York; 2000. p. 179–91. https://doi.org/10.1007/978-1-4612-1284-3_4.
22. Knorr-Held L. Bayesian modelling of inseparable space-time variation in disease risk. *Stat Med.* 2000;19(17–18):2555–67.
23. Lambert, B., *A student's guide to Bayesian statistics.* A Student's Guide to Bayesian Statistics, 2018: p. 1–520.
24. Hoffman MD, Gelman A. The No-U-Turn sampler: adaptively setting path lengths in Hamiltonian Monte Carlo. *J Mach Learn Res.* 2014;15(1):1593–623.
25. Kianfar N, et al. Spatio-temporal modeling of COVID-19 prevalence and mortality using artificial neural network algorithms. *Spatial Spatio-temporal Epidemiol.* 2022;40:100471.
26. Kianfar N, Mesgari MS. GIS-based spatio-temporal analysis and modeling of COVID-19 incidence rates in Europe. *Spatial Spatio-temporal Epidemiol.* 2022;41:100498.
27. Gelman A, Carlin JB, Stern HS, Dunson DB, Vehtari A, Rubin DB. *Bayesian data analysis.* Chapman and Hall/CRC; 2013. <https://doi.org/10.1201/b16018>.
28. Watanabe S, Opper M. Asymptotic equivalence of Bayes cross validation and widely applicable information criterion in singular learning theory. *J Mach Learn Res.* 2010;11(12):3571–94.
29. Cheng F-Y, et al. Using machine learning to predict ICU transfer in hospitalized COVID-19 patients. *J Clin Med.* 2020;9(6):1668.
30. Mangalathu S, et al. Data-driven machine-learning-based seismic failure mode identification of reinforced concrete shear walls. *Eng Struct.* 2020;208:110331.
31. Perera S. *A Bayesian spatial interaction framework for optimal facility location in urban environments.* Warwick: University of Warwick; 2022.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.