Sri Lankan Journal of Applied Statistics, Vol (26-3)

Assessing Policy and Environmental Impacts on Agricultural Yield Trends Using Bayesian Spline and Interrupted Time Series Regression Models

A. W. L. P. Thilan*

Department of Mathematics, Faculty of Science, University of Ruhuna, Sri Lanka

*Corresponding author: pubudu@maths.ruh.ac.lk

Received: 06th June, 2025/ Revised: 22nd September, 2025/ Accepted: 26th September, 2025/
Published: 31st December, 2025

ABSTRACT

Agriculture is a vital sector in Sri Lanka, yet it has been increasingly challenged by environmental disturbances and policy shifts. This study analyzes the impact of two significant disruptions—the 2017 drought and the 2021 fertilizer ban—on paddy yield trends. To address the complexities of these influences, Bayesian spline regression was employed to model non-linear and seasonal yield patterns, and interrupted time series regression was used to capture abrupt changes in response to external shocks. Uncertainty and the causal effects of these interventions were quantified through posterior distributions, estimated using the `brms` *package and the NUTS algorithm. The results offer nuanced insights into how environmental and policy factors shape agricultural productivity, with implications for researchers and policymakers concerned with sustainable agriculture and climate adaptation. Overall, this study underscores the value of Bayesian modeling in assessing multifaceted, real-world issues that span agriculture, environment, and policy.*

Keywords: Bayesian modeling, Environmental disturbances, Seasonal patterns, Non-linear patterns, Sri Lanka agriculture, Multi-factorial influences.

1 Introduction

Agriculture plays a vital role in the economic development and food security of many developing countries, including Sri Lanka (Gunawardena and de Silva , 2021). However, this sector is increasingly challenged by policy changes and environmental disruptions. Recently, Sri Lanka initiated a significant policy shift towards organic farming by implementing a ban on chemical fertilizers starting in 2021 (Jayasuriya and Jayasinghe , 2016). Alongside this policy, severe droughts and other climate-related events have further impacted agricultural productivity. These combined factors have had notable effects on paddy yields, a staple crop critical to national food security and rural livelihoods.

Understanding how policy measures and environmental shocks jointly influence agricultural productivity is crucial for developing resilient and sustainable farming systems (Ruppert and Wand , 2009; Smith and Wang , 2020; Zhao and Li , 2018). Agricultural production does not occur in isolation; it is shaped by a complex set of interacting factors. Policy interventions such as changes in fertilizer regulations or subsidies can directly affect farming practices and input use, thereby influencing crop yields (Jayasuriya and Jayasinghe , 2016; Gunawardena and de Silva , 2021). Meanwhile, environmental disturbances-including droughts, floods, and shifting climate patterns-add further variability and risk to agricultural outputs (Kumar and Singh , 2019). These factors often do not act independently but interact in ways that can amplify or mitigate their individual effects (Lobell and Field , 2011; Wheeler and Von Braun , 2013). Although each has been studied extensively on its own, there is a notable gap in research that simultaneously examines and quantifies the combined and potentially interacting impacts of policy changes and environmental shocks on crop yield trends (Ruppert and Wand , 2009; Gelman and Hill , 2017). Addressing this gap is essential for generating comprehensive insights that can inform adaptive agricultural policies and promote food security under changing environmental conditions (FAO , 2019).

In this study, the effects of the 2021 fertilizer ban and the 2017 drought on paddy yield patterns in Sri Lanka are analyzed using a Bayesian hierarchical modeling framework that integrates spline regression with interrupted time series analysis (Gelman and Hill , 2017; Thilan et al. , 2022). The hierarchical structure allows the model to account for year-to-year variability and seasonal effects, which are important because paddy yields naturally fluctuate across years and between the *Maha* and *Yala* seasons. Bayesian spline regression captures complex, non-linear trends and seasonal fluctuations while providing a full characterization of uncertainty (Kass and Raftery , 1994). Interrupted time series regression complements this approach by detecting abrupt shifts in yields associated with policy or environmental events, enabling a detailed evaluation of their impacts. By combining these methods within a hierarchical framework, the study provides a robust and flexible approach to disentangling gradual trends from sudden changes in agricultural productivity.

Bayesian methods provide a powerful and flexible framework for analyzing complex data, particularly when dealing with uncertainty and non-linear relationships. Unlike traditional approaches, Bayesian inference treats model parameters as random variables, allowing direct quantification of uncertainty through posterior distributions and enabling the incorporation of prior knowledge that can be updated as new data emerge. This probabilistic approach offers enhanced interpretability and predictive accuracy, making it especially suitable for evaluating the effects of policy interventions and environmental disturbances on agricultural productivity. In this study, the models are implemented using the `brms` package and the No-U-Turn Sampler (NUTS) algorithm, which together facilitate efficient Bayesian inference and flexible modeling (Carpenter *et al.* , 2015). This combined approach offers a robust framework for disentangling the dynamic interactions between policy measures, environmental shocks, and paddy yield trends, thereby supporting more informed decision-making in agricultural policy and management.

The primary aim of this study is to quantify the causal impacts of key policy and environmental events on paddy yield trends in Sri Lanka. Specifically, the objectives are to: (i) characterize the non-linear and seasonal patterns of paddy yields; (ii) identify sudden changes linked to policy and environmental disruptions; and (iii) provide actionable insights to support evidence-based agricultural policymaking and sustainable development. This work contributes to a deeper understanding of how multiple factors shape agricultural outcomes, with significant implications for research, policy formulation, and enhancing climate resilience in agriculture.

2 Motivating Example

Rice is the staple food of the country and plays an important role in the diet and culture of Sri Lanka and has also been a key crop for food security and rural livelihood (Weerahewa et al. , 2012). Paddy is cultivated in wetlands across all districts and it is cultivated as two growing seasons; *Maha* and *Yala*. Those seasons correspond with the country's monsoons—the *Maha* season overlaps with the Northeast monsoon (September to March), and the *Yala* season overlaps with the Southwest monsoon (April to August) (Amarasingha et al. , 2015). These seasonal patterns determine the timing of paddy sowing and harvesting. However, paddy cultivation in Sri Lanka has rarely reached its full potential, frequently hindered by water scarcity, unpredictable rainfall, and limited resources (De Silva , 2020).

For this study, data has been sourced from the Agriculture and Environment Statistics Division of the Department of Census and Statistics. This division is responsible for collecting, processing, and sharing agricultural data in Sri Lanka, including statistics on the extent and production of paddy (Department of Census and Statistics , 2024). The data shows significant variability in paddy yields, particularly exogenous examples such as the drop in 2017 and 2018 shown in Figure 1. The sharp decline corresponds mainly to the drought in 2017, during which water usage for irrigation decreased noticeably nationwide on paddy fields throughout Sri Lanka (Wijeratne , 2018). That drought was exacerbated by flooding in some areas, which inflicted further damage to the crops and resulted in the sharp decline in yield (Gunathilaka , 2018). Matters deteriorated further in 2018, when the government's previous policy to limit chemical fertilizer use started to bite. Paddy production had seen a consistent decline due to lack of good farmers, ammonia fertilizer production, which was a necessary input, decreased leading to nutritional deficiencies in soil (Weerakoon , 2021). It is this combination of weather events and policy shifts that partially accounts for the declining production seen in these years.

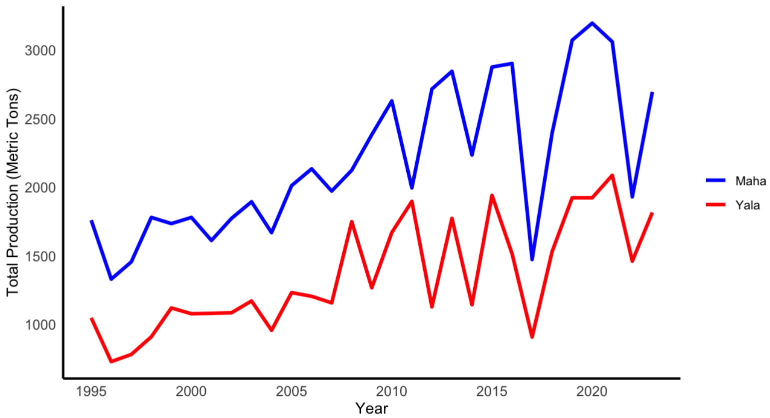

Fig. 1: Fluctuating paddy yield trends across the *Maha* and *Yala* seasons.

Policy interventions in recent years have added another layer of complexity for paddy cultivation. In particular, the Sri Lankan government's 2021 decision on banning chemical fertilizer, which was part of a countrywide "organic farming" move, has greatly reduced paddy yields (Dharmasena , 2021). Although the ban was intended to encourage sustainability, there have been discussions about its immediate and long-term impact on agricultural productivity (Abeywardana , 2022). The impacts are especially seen in the varying yield patterns of the *Maha* and *Yala* seasons (illustrated in Figure 1) which result due to the interaction of socio-economic, environmental, and policy-oriented influences (Marambe , 2022).

This study seeks to model the influence of both policy interventions and climate-related changes on paddy yield trends using Bayesian spline regression and interrupted time-series regression analysis. These methodologies are particularly well-suited to:

1. **Capturing the non-linear and seasonal nature of paddy yield trends**, which are influenced by monsoonal cycles, resource availability, and external shocks.

2. **Identifying structural breaks and shifts in yield patterns** caused by specific policy interventions and climatic disturbances such as the 2021 fertilizer ban and droughts.

3. **Quantifying the impacts of policy measures and climatic factors** in a probabilistic framework, enabling robust inferences about the magnitude and duration of their effects.

By examining the *Maha* and *Yala* seasons in detail, this research aims to provide actionable insights into how Sri Lanka's agricultural policies and climate variability have shaped paddy productivity. Ultimately, it seeks to contribute to evidence-based strategies that balance the country's sustainability goals with the need to ensure food security and economic stability for farming communities.

3 Data and Data Management

The dataset introduced in Section 2 was compiled from the publication *PADDY STATISTICS 2023/2024 MAHA SEASON* issued by the Department of Census and Statistics, Sri Lanka, specifically Table 6: *Harvested Extent, Average Yield, and Production of Paddy by Season (1995–2024)* Department of Census and Statistics (2024). The data are presented in wide format, with each row corresponding to a year (1995–2023) and separate columns for the *Maha* and *Yala* seasons.

Variables taken directly from the reports:

- Gross_Extent_Harvested_Maha / Gross_Extent_Harvested_Yala – total area harvested in the respective season.
- Total_Production_Maha / Total_Production_Yala – total paddy production in the respective season.

Variables coded based on literature:

- Drought_Impact – indicator variable for the 2017 drought (1 for 2017, 0 otherwise).
- Fertilizer_Restriction – indicator variable for fertilizer restrictions (1 from 2018 onwards, 0 before).
- Fertilizer_Ban – indicator variable for the fertilizer ban (1 from 2021 onwards, 0 before).

To facilitate analysis across seasons, the dataset was reshaped into a long format, yielding 58 season-level observations, where each row represents a season-year observation with associated total production and harvested extent. The dataset is complete with no missing values, providing a reliable foundation for subsequent modeling.

4 Methods

The model for predicting paddy production assumes that the observed total production, y_{st}, follows a Normal distribution. This assumption is appropriate due to the continuous nature of paddy production outcomes (e.g., yield measured in kilograms or tons per hectare):

$$y_{st} \sim N(\mu_{st}, \sigma^2),$$

where μ_{st} represents the mean yield for season s and time t, and σ^2 is the variance capturing random fluctuations due to unobserved factors such as measurement errors and unpredictable weather conditions. The assumption of normality is justified by the Central Limit Theorem, as paddy production yields result from a combination of multiple independent effects (rainfall, soil quality, farming practices), which aggregate into a Normal distribution (Gelman et al. , 2013).

The mean response μ_{st} is modeled as:

$$\begin{aligned}\mu_{st} = f(\text{Year}) &+ \beta_1 \text{Drought\_Impact}_t + \beta_2 \text{Fertilizer\_Restriction}_t + \beta_3 \text{Fertilizer\_Ban}_t \\ &+ \beta_4 \text{Seasons}_s + \beta_5 \text{Gross\_Extent\_Harvested}_{st} + u_{\text{Year}_t} + v_{\text{Seasons}_s} + \epsilon_{st}\end{aligned} \tag{1}$$

where:

- $f(\text{Year})$ represents a smooth function of time (Year) modeled using a low-rank thin-plate spline, which accounts for nonlinear effects over time (Wood , 2003).
- $\beta_1, \beta_2, \beta_3, \beta_4, \beta_5$ are the coefficients for the covariates:
  - Drought_Impact_t, $\text{Fertilizer\_Restriction}_t$, and Fertilizer_Ban_t are binary variables indicating the occurrence of these events at time t.
  - Seasons_s is a factor variable indicating the season (e.g., *Maha* or *Yala*).
  - $\text{Gross\_Extent\_Harvested}_{st}$ is a directly observed variable from the dataset, representing the total land area harvested (in hectares) in season s and year t.
- $u_{\text{Year}_t} \sim \mathcal{N}(0, \sigma^2_{\text{Year}})$ is a random intercept for year, capturing unobserved variability between years not explained by the fixed effects or smooth time trend.
- $v_{\text{Seasons}_s} \sim \mathcal{N}(0, \sigma^2_{\text{Season}})$ is a random intercept for season, capturing year-to-year fluctuations within each season and allowing the model to account for unobserved seasonal effects beyond the fixed seasonal difference.
- $\epsilon_{st} \sim \mathcal{N}(0, \sigma^2)$ represents the residual error term.

The smooth function $f(\text{Year})$ is modeled using a thin-plate spline to capture nonlinear temporal effects. This method is chosen for its efficiency, robustness to knot placement, and ability to handle large datasets (Wood , 2003). The thin-plate spline is defined as:

$$f(\text{Year}) = \sum_{k=1}^{K} \delta_k |\text{Year} - \eta_k|^3, \tag{2}$$

where δ_k are random coefficients, and η_k are knots (e.g., quantiles of the covariate Year). K represents the number of knots, which balances model flexibility with computational efficiency (Ruppert, Wand, and Carroll , 2003). To determine the optimal complexity for the spline component, models with values $k = 3, 4, \ldots, 10$ were fitted, where k specifies the number of basis functions that generate the smooth term. The values of k were selected to evaluate flexibility of the model while avoiding overfitting. Model selection was carried out using the Leave-One-Out Information Criterion (LOOIC), which weights model complexity against fit (Vehtari, Gelman, and Gabry , 2017). For each k, the corresponding LOOIC value was computed and compared. Finally, a scree plot was produced to illustrate how model performance changed with varying feature counts k, thereby identifying the optimal model complexity (Spiegelhalter et al. , 2002).

Interventions, such as changes in policies like the introduction of a fertilizer ban, natural disturbances like drought and rainfall events can affect paddy production. Such events can cause abrupt changes in production trends. An interrupted time series (ITS) regression approach can be used to estimate the effect of the fertilizer ban policy and natural disturbances (Linden , 2015). This enables us to capture sudden shifts in the trend induced by interventions occurring at distinct time points. We suspect the fertilizer ban will induce (more or less) a step-change in the mean yield and a gradual change in the slope as time goes on (assuming reduced availability of fertilizers steepens the slope), while drought will have a more abrupt and extreme effect on production levels (Bernal, Cummins, and Gasparrini , 2017).

The regression model can be extended to include the following components, with the ITS approach applied through indicator variables:

- Drought_Impact_t is an indicator variable for drought events, where $\text{Drought\_Impact}_t = 1$ if drought conditions are present at time t and $\text{Drought\_Impact}_t = 0$ otherwise.

- $\text{Fertilizer\_Restriction}_t$ is an indicator variable for fertilizer restriction events, where $\text{Fertilizer\_Restriction}_t = 1$ if fertilizer restrictions are imposed at time t and $\text{Fertilizer\_Restriction}_t = 0$ otherwise.

- Fertilizer_Ban$_t$ is an indicator variable for the fertilizer ban policy, where Fertilizer_Ban$_t = 1$ after the policy intervention and Fertilizer_Ban$_t = 0$ before the policy is implemented.
- $\beta_1, \beta_2, \beta_3$ represent the coefficients that capture the immediate changes in the mean yield due to drought impact, fertilizer restriction, and fertilizer ban interventions, respectively.

ITS can account for the effects of interventions that were planned to set in motion as the intervention is implemented, even if they are not fully implemented (Wagner et al. , 2002; Thilan , 2025). This model situates the effects of drought, fertilizer restrictions, and the fertilizer ban by adopting indicator variables, on the understanding that the joint temporal impacts and the gradual changes, from these to production trends, are contributors of contemporary relevance. This approach offers a strong basis for assessing the aggregate effect of policy changes and natural disturbances on paddy production, allowing for a separation of intervention effects from preexisting trends or external factors (Shadish, Cook, and Campbell , 2015).

By including a fixed effect for Seasons$_s$, the model directly captures seasonal differences under a single model, measuring differences between the *Yala* and the *Maha* phases. This enables simultaneous evaluation of covariates, including Drought_Impact$_t$, Fertilizer_Restriction$_t$, Fertilizer_Ban$_t$, and Gross_Extent_Harvested$_{st}$ for both respective seasons, improving efficiency and interpretability of models (Gelman et al. , 2013). Incorporating Seasons$_s$ directly into the analysis ensures that systematic differences between seasons are accounted for, while production trends can still be directly compared. Fitting a single model with all these factors significantly improves the accuracy of estimates and enhances the interpretability of the interplay of seasonality with the other key drivers of μ_{st} (Wood , 2017). The combined model makes it possible to assess paddy production trends that include both direct and longer-term effects associated with the fertilizer ban intervention, while also recognizing the unique characteristics for each *Yala* and *Maha* season as well as filtering out unobserved heterogeneity through random effects (Ruppert, Wand, and Carroll , 2003). This approach aligns with best practices in hierarchical modeling, where fixed and random effects are used to account for both observed and unobserved sources of variation (Gelman et al. , 2013; Thilan et al. , 2024).

Bayesian modeling estimates the posterior distribution of model parameters by combining observed data with prior distributions through Bayes' Theorem (Gelman et al. , 2013). Unlike frequentist methods, which yield point estimates and confidence intervals, Bayesian methods provide entire posterior distributions that directly quantify uncertainty (Kruschke , 2014; Thilan et al. , 2023). This approach is particularly effective for models with complex hierarchical structures, small sample sizes, and random effects (Gelman , 2006). It also allows for the inclusion of prior knowledge or domain expertise, leading to more intuitive uncertainty quantification (McElreath , 2020).

Furthermore, the use of Markov Chain Monte Carlo (MCMC) sampling enables robust inferences even for complex, non-standard models (Betancourt , 2017). Overall, these advantages make Bayesian modeling well-suited for evaluating policy and environmental impacts on agricultural yield trends.

The `brms` package was chosen for Bayesian modeling because it offers flexible and user-friendly syntax for specifying a wide range of models in `R` (Bürkner , 2017). Built on the `Stan` probabilistic programming language, `brms` enables robust posterior sampling with a familiar formula-based interface, similar to `lm` and `lme4` (Stan Development Team , 2023). This makes it easier for researchers to transition from frequentist to Bayesian methods. The package supports random effects, smooth terms, and a variety of models, including generalized linear models (GLMs), generalized additive models (GAMs), hierarchical models, and non-linear models, all within a single framework (Bürkner , 2018). It also offers essential posterior diagnostics, such as posterior predictive checks and convergence diagnostics, enhancing the reliability of model inferences (Bürkner , 2017). Integration with `Stan` ensures efficient computation and scalability, making `brms` well-suited for analyzing complex, hierarchical data while incorporating uncertainty and prior information. Details of the model implementation are provided in Section 9.

Default weakly informative priors were used for all fixed effects and intercepts to provide regularization while allowing the data to largely inform the posterior estimates. For the spline term, `s(Year, bs = "tp")`, smoothness was estimated from the data with default priors, enabling the model to flexibly capture nonlinear temporal trends. The random effects for Year and Season were assigned default hierarchical priors, reflecting uncertainty in year-to-year and seasonal variability. Hyperparameters controlling the MCMC sampling—such as `iter = 10000`, `warmup = 3000`, `adapt_delta = 0.9999`, and `max_treedepth = 15`—were chosen to ensure stable convergence and avoid divergent transitions during sampling. These settings collectively allow the model to produce robust posterior estimates while accommodating the hierarchical structure and complex dependencies present in the agricultural yield data.

5 Results

Using a systematic approach, the optimal spline complexity (k) was identified for each version of the smooth function $s(\mathit{Cumulative\_Month}_i, \mathit{Phase}_i)$ in the model as per Equation 1. The scree plot in Figure 2 depicts the pattern of LOOIC values based on spline complexity (k) having the LOOIC decreasing from $k = 1$ through $k = 4$ with $k = 4$ having the lowest LOOIC value indicating the best fit. After $k = 4$, the values of LOOIC start to increase, indicating overfitting. Then the squared values stabilize again at higher complexities with negligible improvement. Hence, $k = 4$ was the final choice, achieving a proper trade-off between flexibility and performance of the model.

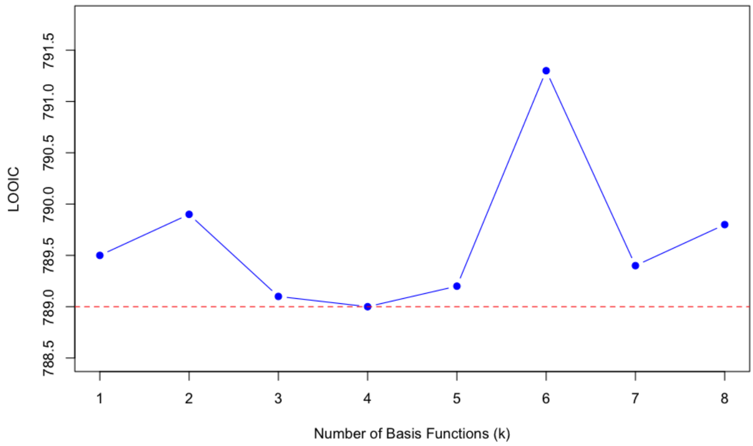

Fig. 2: Scree plot of LOOIC values as a function of spline complexity (k).

A summary of the model results is presented in Table 1, highlighting the predictors influencing *Total_Production*. The smoothing spline parameter *sds(sYear_1)* has an estimate of 286.47 (95% CI: [8.88, 982.21]), indicating moderate year-to-year variability in production. The random effects for *Season* (*sd(Intercept)*: 555.73, 95% CI: [18.84, 1964.68]) and *Year* (*sd(Intercept)*: 125.16, 95% CI: [58.56, 193.26]) suggest that seasonal and annual differences contribute meaningfully to model fit, with seasonal effects being more pronounced.

Among the regression coefficients, *Fertilizer Ban* shows a strong negative effect (-533.79, 95% CI: [-836.03, -206.70]), indicating that banning fertilizers substantially reduces total production. *Fertilizer Restriction* has a positive estimate (174.45), although its confidence interval spans zero, reflecting uncertainty about its impact. *Drought Impact* has a small negative estimate (-81.23) with a wide interval including zero, suggesting a less clear effect in this model. *Gross Extent Harvested* exhibits a consistent positive association with production (3.80, 95% CI: [3.01, 4.57]), confirming its significant role. The effect of *Season (Yala)* is positive but highly uncertain (109.61, 95% CI: [-2188.79, 2412.70]). The smoothing spline for year, *sYear_1*, indicates additional temporal variability, though with a wide interval. The residual standard deviation (*Sigma*) is 124.70 (95% CI: [94.83, 167.77]), reflecting unexplained variation in total production. Overall, these results underscore the importance of land area harvested and the fertilizer ban policy, while the effects of drought and seasonality remain less definitive.

Figure 3 illustrates the observed and predicted trends in paddy production for the *Maha* and *Yala* seasons. The *Maha* season shows consistently higher production levels with a steady upward trend over time, whereas the *Yala* season reflects lower production with greater year-to-year variability.

Table 1: Summary of Regression Coefficients and Smoothing Splines

Category	Parameter	Estimate	Est. Error	*l-95% CI*	*u-95% CI*	Rhat	Bulk ESS	Tail ESS
Smoothing Spline Hyperparameters								
	sds(*sYear*_1)	286.47	264.63	8.88	982.21	1.00	8883	11750
Multilevel Hyperparameters								
Season (2 levels)	sd(*Intercept*)	555.73	536.31	18.84	1964.68	1.00	11980	13074
Year (29 levels)	sd(*Intercept*)	125.16	33.56	58.56	193.26	1.00	5923	5954
Regression Coefficients								
	Intercept	-82.39	687.36	-1512.80	1337.15	1.00	12746	11459
	Drought Impact	-81.23	214.16	-503.10	350.32	1.00	14303	17677
	Fertilizer Restriction	174.45	162.51	-140.85	513.00	1.00	12093	13887
	Fertilizer Ban	-533.79	156.94	-836.03	-206.70	1.00	12322	13985
	Season (Yala)	109.61	1085.69	-2188.79	2412.70	1.00	12344	10550
	Gross Extent Harvested	3.80	0.40	3.01	4.57	1.00	19620	20274
	*sYear*_1	519.83	988.87	-1966.22	2173.13	1.00	9500	8606
Further Distributional Parameters								
	Sigma	124.70	18.72	94.83	167.77	1.00	7851	9234

A notable decline is observed around 2017–2018, corresponding to adverse weather conditions and policy changes, after which production levels recover. The prediction lines closely follow the observed data, and the shaded 95% prediction intervals remain relatively narrow, indicating that the model captures the production dynamics effectively. These results demonstrate distinct seasonal patterns, with *Maha* contributing more substantially and more stably to overall paddy production compared to *Yala*.

Posterior predictive samples from the fitted model were generated to assess its predictive ability. The Figure 4 displays the distribution of predicted outcomes, representing the expected value based on the posterior parameters from the model. These predictions offer insight into the model's performance by capturing the underlying trends and variability in the data. The plotted graph demonstrates the fit of the model by comparing it to the collected data, highlighting how well the model aligns with the actual patterns. Overall, the model adequately captures the fundamental patterns in the data, with the predicted outcomes aligning closely with their observed counterparts, indicating an acceptable fit.

To further assess model adequacy, residual diagnostics and posterior predictive checks were conducted to verify that the model captured the underlying data structure without leaving systematic patterns unaccounted for. The partial autocorrelation function (PACF; Figure 8) showed no substantial temporal dependence beyond lag 1, and the Ljung-Box test was non-significant ($p = 0.1865$), indicating no residual autocorrelation. Observed values aligned well with predicted outcomes, as illustrated by the posterior predictive scatterplot (Figure 9), and detailed results are provided in the Supplementary Material (Section 9.9). Overall, these evaluations confirm that the model reliably represents the key data patterns and variability without systematic deviations.

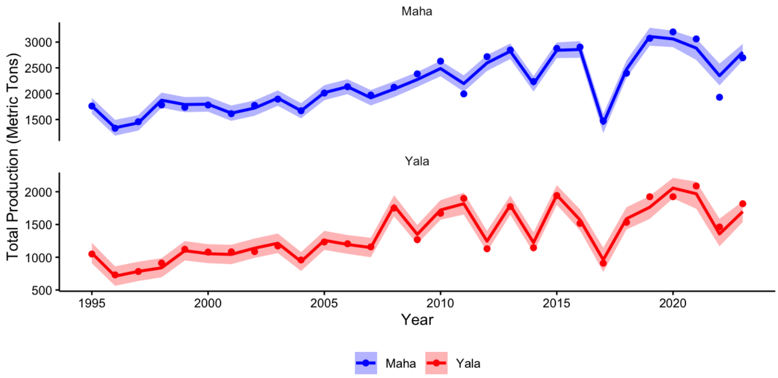

Fig. 3: Trends in total paddy production during Sri Lanka's *Maha* (blue) and *Yala* (red) seasons from 1995 onwards. Points represent observed data, solid lines indicate predicted values, and shaded regions represent 95% prediction intervals, highlighting variability and uncertainty. Key declines in 2017–2018 reflect environmental and policy impacts.

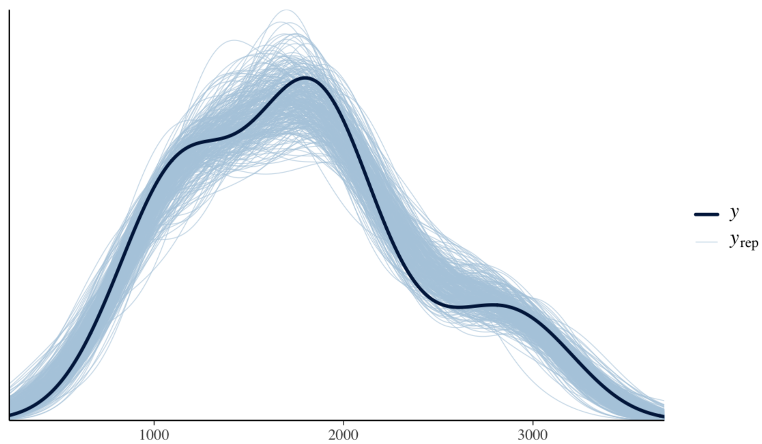

Fig. 4: Density Overlay of Observed vs Predicted Production.

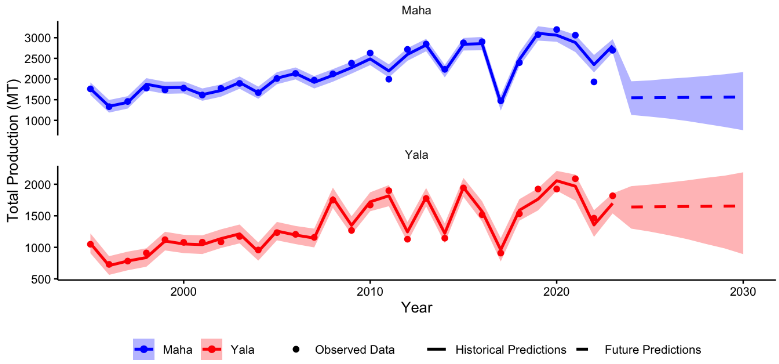

Fig. 5: Predicted vs Observed Total Production (Historical and Future).

The future projections for total production in the *Maha* and *Yala* seasons, shown in Figure 5, exhibit distinct seasonal patterns. The *Maha* season is projected to continue growing at a slower rate, with a relatively narrower confidence band, indicating higher confidence in its estimates. In contrast, the *Yala* season shows continued growth with a wider confidence band, reflecting greater uncertainty in its future outcomes. These differences in uncertainty highlight seasonal variations in production predictability, with *Yala* showing greater variability in future trends. The methodology for generating these projections, including dataset extension and visualization of historical versus future predictions, is detailed in the Supplementary Material (Section 9.10).

6 Discussion

Bayesian spline regression and interrupted time series analysis were employed to assess the impact of the Sri Lankan government's fertilizer ban imposed in early 2021, along with precursor restrictions and environmental factors such as drought, on paddy yield trends. Beyond immediate policy implications, the study demonstrates the practical application of advanced statistical techniques in real-world agricultural and policy contexts, contributing to better-informed policy evaluation and decision-making.

A key finding is the nuanced impact of fertilizer policies: the fertilizer restriction exhibited a weak positive effect on paddy production with considerable uncertainty, whereas the fertilizer ban corresponded with a pronounced negative effect, as indicated by the credible interval excluding zero. Some parameters, including the drought impact and spline term, have relatively wide credible intervals, reflecting high uncertainty that may arise from limited data or collinearity among predictors. This uncertainty should be considered when interpreting the results. These results highlight the need for policymakers to interpret such interventions with caution, recognizing that effects may not be uniform or immediate. Continued monitoring and research remain essential to fully understand long-term and indirect consequences.

Substantial variability in paddy production across years and seasons was captured by the spline term (*sds(sYear_1)*) and the random effects for Season and Year. The wide credible intervals for these terms indicate considerable uncertainty in estimating temporal trends, which is likely influenced by the relatively small number of season-level observations and correlation among predictors. Including both Year and Season as random effects allows the model to account for the hierarchical structure in the data while quantifying uncertainty appropriately. This underscores the importance of incorporating long-term temporal factors and seasonal patterns in agricultural analyses. Ignoring these dynamic influences risks misleading estimates of policy and environmental impacts and may reduce the effectiveness of interventions.

Reductions in paddy yield arise from the interplay of policy measures and environmental conditions, including droughts and other stressors. Bayesian spline regression provided the flexibility to disentangle these overlapping effects, demonstrating the value of advanced statistical models in complex, real-world settings. The uncertainty in some parameter estimates highlights the need to interpret results cautiously, acknowledging that the magnitude of certain effects may vary. This integrated approach serves as a case study for applying quantitative tools to support welfare and agricultural policy analysis.

A major strength and significant contribution of this study is the adoption of a probabilistic Bayesian framework, which explicitly quantifies uncertainty in estimating the effects of fertilizer policies and drought on paddy production. Unlike traditional frequentist methods that provide single point estimates, our approach offers a range of plausible outcomes. This probabilistic perspective is especially important in agriculture and environmental policy, where decisions have wide-reaching and potentially irreversible consequences. Moreover, our emphasis on transparent reporting through posterior intervals and diagnostic checks can help reduce overconfidence and improve the quality of policy evaluations.

Finally, the broader relevance of our methodological approach extends beyond agriculture. Bayesian spline regression combined with interrupted time series analysis is a versatile toolkit capable of addressing complex, nonlinear, and time-dependent phenomena across sectors-from health care to education to economics. For example, these methods could assess the impacts of health care policy reforms on patient outcomes, evaluate education policy effectiveness on student performance, or analyze the economic consequences of fiscal measures, underscoring the wide utility of Bayesian frameworks in policy research.

7 Conclusion

Our findings underscore the critical role of Bayesian statistical techniques in evaluating the complexities of agricultural policies and environmental sustainability. By leveraging Bayesian spline regression and interrupted time series analysis, this study effectively modeled the nonlinear trends and sudden shifts in paddy yield data-patterns often encountered in real-world agricultural systems. The analysis of fertilizer bans, policy restrictions, and droughts provides nuanced insights into how both policy interventions and environmental variability shape agricultural outcomes over time.

This research also demonstrates the broader applicability of Bayesian methods for addressing complex, multi-factorial challenges across various disciplines. The integration of Bayesian spline regression and interrupted time series analysis offers practical strategies for policymakers, researchers, and analysts in fields such as health, economics, and education. By explicitly quantifying uncertainty and capturing both gradual and abrupt changes, this study highlights how these advanced techniques can support data-informed decision-making in a wide range of real-world applications.

8 Acknowledgment of AI Assistance

This manuscript benefited from the use of artificial intelligence tools. OpenAI's ChatGPT was used to assist with language refinement, improving clarity, and rephrasing of certain passages. All final content was reviewed and approved by the author.

9 SUPPLEMENTARY MATERIAL

9.1 Introduction This appendix provides a detailed description of the R code used to model the paddy production data, incorporating disturbance variables, reshaping the dataset for analysis, exploratory data analysis (EDA), and fitting a Bayesian regression model using the brms package. The code also includes visualization and model diagnostics.

9.2 Required Packages The script starts by installing and loading the required R packages: tidyr, ggplot2, brms, and dplyr for data manipulation, visualization, and Bayesian modeling.

```r
# Install and load required package
# install.packages("tidyr")
library(tidyr)
library(ggplot2)
library(brms)
library(dplyr)
```

Listing 1: Loading Required Libraries

9.3 Data Loading and Disturbance Variables The dataset paddy_season_data.csv is loaded into the environment. Additional disturbance variables are added to the dataset:

- Drought_Impact: Indicates whether the year is 2017, the year affected by drought.
- Fertilizer_Restriction: Indicates whether fertilizer restrictions are in place starting from 2018.
- Fertilizer_Ban: Indicates whether a fertilizer ban is in place starting from 2021.

```r
# Load the dataset
dataset <- read.csv("paddy_season_data.csv")

# Add disturbance variables
# Drought impact in 2017
dataset$Drought_Impact <- ifelse(dataset$Year == 2017, 1, 0)
# Fertilizer restrictions starting from 2018
dataset$Fertilizer_Restriction <- ifelse(dataset$Year >= 2018, 1, 0)
# Fertilizer ban in 2021
dataset$Fertilizer_Ban <- ifelse(dataset$Year >= 2021, 1, 0)
```

Listing 2: Adding Disturbance Variables

9.4 Reshaping the Dataset To prepare the data for modeling, the dataset is reshaped from a wide to a long format using the `pivot_longer` function. The production data for both seasons (*Maha* and *Yala*) is combined into a single column. Additional transformations ensure the correct mapping of variables such as `Season` and `Gross_Extent_Harvested`.

```
# Reshape the dataset for combined modeling
combined_dataset <- dataset %>%
  select(Year, Drought_Impact, Fertilizer_Restriction, Fertilizer_Ban,
         Total_Production_Maha, Total_Production_Yala,
         Gross_Extent_Harvested_Maha, Gross_Extent_Harvested_Yala) %>%
  pivot_longer(
    cols = starts_with("Total_Production"),
    names_to = "Season",
    values_to = "Total_Production"
  ) %>%
  mutate(
    Season = ifelse(Season == "Total_Production_Maha", "Maha", "Yala")
      ,
    Gross_Extent_Harvested = ifelse(Season == "Maha", Gross_Extent_
        Harvested_Maha, Gross_Extent_Harvested_Yala)
  ) %>%
  select(Year, Drought_Impact, Fertilizer_Restriction, Fertilizer_Ban,
         Season, Total_Production, Gross_Extent_Harvested)
```

Listing 3: Reshaping the Dataset for Combined Modeling

9.5 Exploratory Data Analysis (EDA) An EDA was performed to summarize and visualize the data before modeling. This included descriptive statistics and visualizations of relationships between predictors and the response variable. This step helps understand data patterns and detect potential issues such as multicollinearity.

9.5.1 Summary Statistics

To gain a clearer understanding of the dataset, basic summary statistics were computed by season. These statistics provide insights into the central tendency and variability of production and harvested area.

```
# Summary statistics for EDA
eda_summary <- combined_dataset %>%
  group_by(Season) %>%
  summarise(
    Mean_Production = mean(Total_Production),
    SD_Production = sd(Total_Production),
```

```
    Min_Production = min(Total_Production),
    Max_Production = max(Total_Production),
    Mean_Harvested = mean(Gross_Extent_Harvested),
    SD_Harvested = sd(Gross_Extent_Harvested)
  )
print(eda_summary)
```

Listing 4: Summary Statistics by Season

It is evident from Table 2 that the *Maha* season has higher average production and harvested area compared to the *Yala* season. The differences in harvested area are particularly pronounced, with the *Maha* season covering nearly twice the mean area of the *Yala* season.

Table 2: Summary Statistics of Paddy Production and Harvested Area by Season

Season	Mean Production (Metric Tons)	SD Production (Metric Tons)	Min Production (Metric Tons)	Max Production (Metric Tons)	Mean Harvested (Hectares)	SD Harvested (Hectares)
Maha	2188.0	546.0	1331	3197	601.0	115.0
Yala	1356.0	396.0	730	2088	358.0	84.6

9.5.2 Visualizing the Trends in Paddy Production

Trends in total production over the years for both seasons (*Maha* and *Yala*) were visualized as part of EDA. The output of the exploratory scatterplot illustrating the relationship between total production and gross extent harvested for the *Maha* and *Yala* seasons is included in the Introduction (Figure 1). This visualization provides context on production patterns and seasonal differences prior to modeling.

```
# Scatterplot for EDA
ggplot(combined_dataset, aes(x = Gross_Extent_Harvested, y = Total_
    Production, color = Season)) +
  geom_point(size = 2) +
  geom_smooth(method = "lm", se = FALSE) +
  labs(x = "Gross␣Extent␣Harvested", y = "Total␣Production␣(Metric␣
      Tons)") +
  theme_minimal()
```

Listing 5: Scatterplot: Total Production vs Gross Extent Harvested

9.5.3 Visualizing Relationships in Paddy Production

As part of the EDA, relationships between paddy production and key explanatory factors were examined. In particular, the association between total production and the gross extent harvested was explored, alongside the potential impacts of droughts and fertilizer restrictions.

```
# Scatterplot for EDA
ggplot(combined_dataset, aes(x = Gross_Extent_Harvested, y = Total_
    Production, color = Season)) +
  geom_point(size = 2) +
  geom_smooth(method = "lm", se = FALSE) +
  labs(x = "Gross Extent Harvested", y = "Total Production (Metric
      Tons)") +
  theme_minimal() +
  theme(
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line = element_line(color = "black")
  )
# Save plot for LaTeX inclusion
ggsave("figures/yield_trends.png", width = 7, height = 5, dpi = 300)
```

Listing 6: Scatterplot: Total Production vs Gross Extent Harvested

The exploratory scatterplot in Figure 6 illustrates the relationship between total production and harvested extent across the *Maha* and *Yala* seasons. While both seasons show a clear positive association between harvested extent and production, the slopes of the fitted trend lines indicate differences in productivity patterns. The *Maha* season exhibits higher total production levels for larger extents harvested, with points more widely dispersed, reflecting greater variability in yields. In contrast, the *Yala* season shows a tighter clustering of observations and a relatively lower production range, suggesting more consistent but smaller-scale harvest outcomes. This comparison provides valuable context on seasonal production dynamics before modeling.

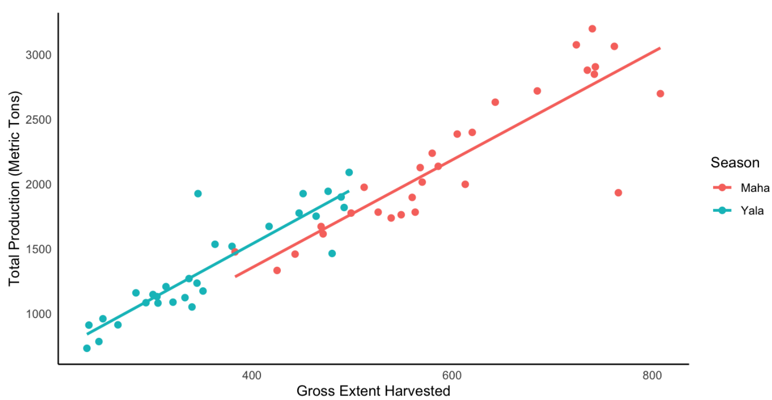

Fig. 6: Scatterplot of total production versus gross extent harvested for the *Maha* and *Yala* seasons.

9.5.4 Visualizing the Impact of Policy Interventions on Paddy Production

The effect of policy interventions, specifically fertilizer restrictions, on total paddy production was examined using boxplots. Such visualizations help reveal differences in production distributions across policy scenarios and seasons, providing preliminary evidence of hierarchical patterns in the data. Similar exploratory visualizations were also considered for other key predictors, such as drought impacts and gross extent harvested, to assess variability and grouping structures prior to model formulation.

```
# Boxplot for EDA
ggplot(combined_dataset, aes(x = factor(Fertilizer_Restriction), y =
    Total_Production, fill = Season)) +
  geom_boxplot() +
  labs(x = "Fertilizer␣Restriction␣(0=No,␣1=Yes)", y = "Total␣
      Production␣(Metric␣Tons)") +
  theme_minimal() +
  theme(
    panel.grid.major = element_blank(),
    panel.grid.minor = element_blank(),
    axis.line = element_line(color = "black")
  )
# Save plot for LaTeX inclusion
ggsave("figures/fertilizer_boxplot.png", width = 7, height = 5, dpi =
    300)
```

Listing 7: Boxplot: Total Production by Fertilizer Restriction

The boxplot in Figure 7 illustrates the distribution of total production for cases with and without fertilizer restrictions, separated by season (*Maha* and *Yala*). It is evident that the *Maha* season generally has higher production values than the *Yala* season across both policy conditions. Additionally, production variability appears higher when fertilizer restrictions are applied, particularly for the *Maha* season. These patterns highlight how policy interventions interact with seasonal differences in production and provide context for the hierarchical modeling of the data.

9.6 Bayesian Model Fitting

The Bayesian model is defined and fitted using the `brms` package. The model includes a thin plate spline for the year variable (`s(Year, bs = "tp")`), disturbance variables, and random effects for `Year` and `Season`. The model is fit using 10,000 iterations with 3,000 warm-up iterations.

```
combined_model <- brm(
  formula = Total_Production ~ s(Year, bs = "tp") + Drought_Impact +
      Fertilizer_Restriction + Fertilizer_Ban + Season +
```

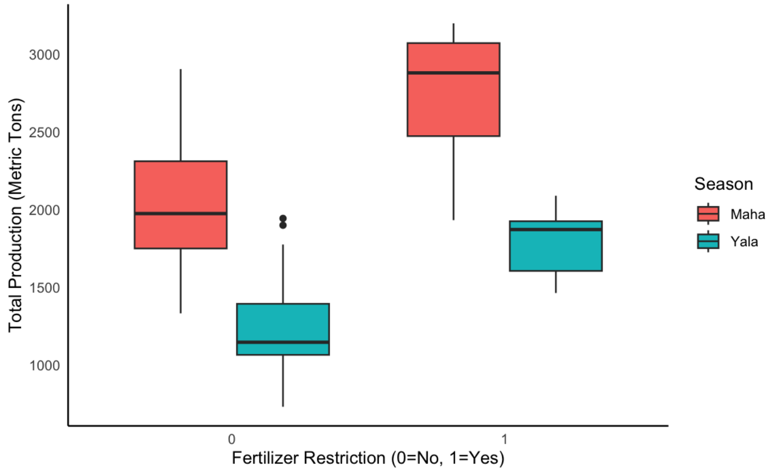

Fig. 7: Boxplot of total paddy production by fertilizer restriction (0 = No, 1 = Yes) for the *Maha* and *Yala* seasons.

```
    Gross_Extent_Harvested +
    (1 | Year) + (1 | Season),
  data = combined_dataset,
  family = gaussian(),
  iter = 10000,
  warmup = 3000,
  control = list(adapt_delta = 0.9999, max_treedepth = 15)
)
```

Listing 8: Fitting the Bayesian Model

9.7 Model Diagnostics and Visualization

Posterior predictive checks were performed using the `pp_check` function to evaluate the model's fit (Figure 4). Residuals from the model were also calculated and visualized to assess model performance.

```
# Posterior predictive check (density overlay) with 300 draws
pp_check_plot_combined <- pp_check(combined_model, type = "dens_
    overlay", ndraws = 300) +
  ggtitle("Posterior␣Predictive␣Check:␣Combined␣Production␣-␣
      Density␣Overlay")
print(pp_check_plot_combined)

# Generate posterior predictive samples
```

```
posterior_samples <- posterior_predict(combined_model, draws =
    100)

# Calculate predicted values (mean across draws) and residuals
combined_dataset$predicted_Total_Production <- apply(posterior_
    samples, 2, mean)
combined_dataset$residuals_combined <- combined_dataset$Total_
    Production - combined_dataset$predicted_Total_Production

# Calculate 95\% prediction intervals
combined_dataset$lower_bound <- apply(posterior_samples, 2,
    function(x) quantile(x, 0.025))
combined_dataset$upper_bound <- apply(posterior_samples, 2,
    function(x) quantile(x, 0.975))

# Visualize residuals
ggplot(combined_dataset, aes(x = residuals_combined)) +
  geom_histogram(bins = 30, fill = "lightblue", color = "black")
    +
  ggtitle("Residuals from Posterior Predictive Check - Combined 
    Model") +
  xlab("Residuals") + ylab("Frequency")
```

Listing 9: Posterior Predictive Check and Residuals

9.8 Prediction and Visualization of Observed vs Predicted Production

The script generates posterior predictive samples and calculates the 95% prediction intervals. Observed and predicted production are visualized with prediction intervals (Figure 3).

```
# Plot observed vs predicted production with prediction intervals
ggplot(combined_dataset, aes(x = Year)) +
  # Observed data as points
  geom_point(aes(y = Total_Production, color = Season, shape = "
    Observed Data"), size = 2) +

  # Predicted values as lines
  geom_line(aes(y = predicted_Total_Production, color = Season),
    size = 1) +
```

```
# Prediction intervals as ribbons
geom_ribbon(aes(ymin = lower_bound, ymax = upper_bound, fill = Season), alpha = 0.3) +

labs(x = "Year", y = "Total␣Production␣(Metric␣Tons)", title = "") +
scale_color_manual(values = c("Yala" = "red", "Maha" = "blue")) +
scale_fill_manual(values = c("Yala" = "red", "Maha" = "blue")) +
scale_shape_manual(name = "", values = c("Observed␣Data" = 16)) +

theme_minimal() +
theme(legend.title = element_blank(),
      legend.position = "bottom",
      panel.grid = element_blank(),
      axis.line = element_line(colour = "black"),
      axis.ticks = element_line(colour = "black"),
      axis.text = element_text(colour = "black")) +
facet_wrap(~ Season, scales = "free_y", nrow = 2)
```

Listing 10: Observed vs Predicted Total Production with Prediction Intervals

9.9 Residual and Predictive Checks To evaluate the adequacy of the fitted model, residual partial autocorrelation, the Ljung-Box test, and posterior predictive checks were examined.

The partial autocorrelation function (PACF) of residuals was computed and plotted using:

```
pacf(residuals_combined, main = "PACF␣of␣Model␣Residuals")
```

Listing 11: PACF of model residuals

The PACF indicated no substantial temporal dependence beyond lag 1, suggesting that the spline term and random effect for Year sufficiently captured temporal variation (Figure 8). In particular, the PACF allows assessment of correlation at a given lag after accounting for correlations at shorter lags, providing a clearer picture of any remaining temporal structure than the standard ACF.

Residual autocorrelation was further assessed with the Ljung-Box test:

```
Box.test(residuals_combined, lag = 20, type = "Ljung-Box")
```

Listing 12: Ljung-Box test for residual autocorrelation

The test yielded a p-value of 0.1865, indicating no significant residual autocorrelation and supporting the conclusion that temporal dependence was adequately addressed by the model.

Posterior predictive checks were performed using:

```
pp_check(combined_model, ndraws = 100)         # density overlay
pp_check(combined_model, type = "scatter_avg") # average predictive
    scatterplot
```

Listing 13: Posterior predictive checks

The density overlay shows close agreement between the observed and simulated data distributions (Figure 4), while the average predictive scatterplot demonstrates good alignment between observed and predicted values (Figure 9). Together, these diagnostics indicate that the model assumptions were reasonably satisfied, and no further time-series adjustments were required.

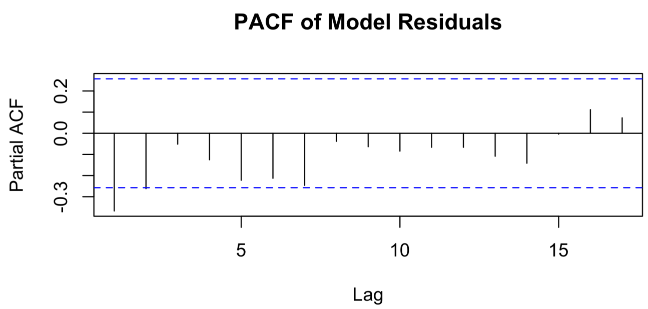

Fig. 8: Partial autocorrelation function (PACF) of residuals from the fitted model. No strong temporal autocorrelation was observed beyond lag 1.

9.10 Future Predictions To visualize future production (2024–2030), the dataset is extended, predictions are generated, and historical vs future predictions are plotted with dashed lines for future values (Figure 5).

```
# Future data preparation
future_years <- 2024:2030
future_data <- expand.grid(
  Year = future_years,
  Season = c("Maha", "Yala"),
  Drought_Impact = 0,
```

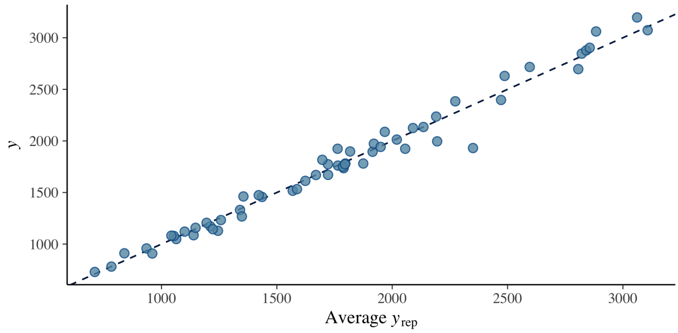

Fig. 9: Posterior predictive scatterplot (average) indicating close alignment between observed and predicted values.

```
  Fertilizer_Restriction = 1,
  Fertilizer_Ban = 1,
  Gross_Extent_Harvested = mean(combined_dataset$Gross_Extent_
      Harvested, na.rm = TRUE)
)

# Generate predictions for future data
future_predictions <- posterior_epred(combined_model, newdata =
    future_data, allow_new_levels = TRUE)
future_data <- future_data %>%
  mutate(predicted_Total_Production = apply(future_predictions,
      2, mean),
         lower_bound = apply(future_predictions, 2, quantile,
             0.025),
         upper_bound = apply(future_predictions, 2, quantile,
             0.975),
         Total_Production = NA,
         is_future = TRUE)

# Combine historical and future data
all_data <- rbind(
  combined_dataset %>% select(Year, Season, Total_Production,
      predicted_Total_Production, lower_bound, upper_bound) %>%
      mutate(is_future = FALSE),
```

```
  future_data %>% select(Year, Season, Total_Production,
      predicted_Total_Production, lower_bound, upper_bound, is_
      future)
)

# Plot combined historical and future predictions
ggplot(all_data, aes(x = Year)) +
  geom_point(aes(y = Total_Production, color = Season, shape = "
      Observed Data"), size = 2) +
  geom_line(aes(y = predicted_Total_Production, color = Season,
      linetype = is_future), size = 1) +
  geom_ribbon(aes(ymin = lower_bound, ymax = upper_bound, fill =
      Season), alpha = 0.3, color = NA) +
  labs(x = "Year", y = "Total Production (Metric Tons)", title =
      "") +
  scale_color_manual(values = c("Yala" = "red", "Maha" = "blue"))
       +
  scale_fill_manual(values = c("Yala" = "red", "Maha" = "blue"))
      +
  scale_shape_manual(name = "", values = c("Observed Data" = 16))
       +
  scale_linetype_manual(name = "", values = c("FALSE" = "solid",
      "TRUE" = "dashed"),
                        labels = c("Historical Predictions", "Future
                            Predictions")) +
  theme_minimal() +
  theme(legend.title = element_blank(),
        legend.position = "bottom",
        panel.grid = element_blank(),
        axis.line = element_line(colour = "black"),
        axis.ticks = element_line(colour = "black"),
        axis.text = element_text(colour = "black")) +
  facet_wrap(~ Season, scales = "free_y", nrow = 2)
```

Listing 14: Historical and Future Predictions with 95% Prediction Intervals

References

Abeywardana, N. (2022). Long-Term Impacts of Organic Farming Policies in Sri Lanka. *Agricultural Policy Journal*, 14(2), 67–82.

Amarasingha, R. P. K., Rathnayake, R. K., & Suriyagoda, S. (2015). Seasonal Patterns of Paddy Cultivation in Sri Lanka. *Journal of Agricultural Science*, 15(3), 78–92.

Angrist, J. D., & Pischke, J.-S. (2009). *Mostly Harmless Econometrics: An Empiricist's Companion*. Princeton University Press.

Bernal, J. L., Cummins, S., & Gasparrini, A. (2017). Interrupted Time Series Regression for the Evaluation of Public Health Interventions: A Tutorial. *International Journal of Epidemiology*, 46(1), 348–355.

Betancourt, M. (2017). A Conceptual Introduction to Hamiltonian Monte Carlo. arXiv:1701.02434 [stat.ME]. Retrieved from `https://arxiv.org/abs/1701.02434`.

Bürkner, P.-C. (2017). brms: An R Package for Bayesian Multilevel Models Using Stan. *Journal of Statistical Software*, 80(1), 1–28.

Bürkner, P.-C. (2018). Advanced Bayesian Multilevel Modeling with the R Package brms. *The R Journal*, 10(1), 395–411.

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., & Goodrich, B. (2015). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1).

Carpenter, B., Gelman, A., Hoffman, M. D., Lee, D., Goodrich, B., Betancourt, M., Brubaker, M., Guo, J., Li, P., & Riddell, A. (2017). Stan: A Probabilistic Programming Language. *Journal of Statistical Software*, 76(1), 1–32.

De Silva, S. (2020). Challenges in Sri Lankan Paddy Cultivation. *Sri Lankan Journal of Agriculture*, 12(4), 34–45.

Department of Census and Statistics. (2024). *PADDY STATISTICS 2023/2024 MAHA SEASON*. Colombo: Department of Census and Statistics, Sri Lanka. Retrieved from `https://www.statistics.gov.lk/Agriculture/StaticalInformation/PaddyStatistics/PaddyStatistics/2023_2024_MahaSeason`.

Department of Census and Statistics, Sri Lanka. (2024). Paddy Statistics. Agriculture and Environment Statistics Division. Available: `https://www.statistics.gov.lk/Agriculture/StaticalInformation/PaddyStatistics` (accessed 2025-03-13).

Dharmasena, P. (2021). Organic Farming Policy and Its Implications for Sri Lanka. *Journal of Sustainable Agriculture*, 25(3), 40–55.

Eilers, P. H. C., & Marx, B. D. (1996). Flexible Smoothing with B-splines and Penalties. *Statistical Science*, 11(2), 89–102.

Food and Agriculture Organization (FAO). (2019). *The State of Food and Agriculture 2019: Moving Forward on Food Loss and Waste Reduction*. FAO.

Gelman, A. (2006). Prior Distributions for Variance Parameters in Hierarchical Models. *Bayesian Analysis*, 1(3), 515–534.

Gelman, A., & Hill, J. (2017). *Data Analysis Using Regression and Multilevel/Hierarchical Models*. Cambridge University Press.

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian Data Analysis* (3rd ed.). Chapman and Hall/CRC.

Gunathilaka, R. P. D. (2018). Flood Impacts on Paddy Yields in Sri Lanka. *Journal of Water and Climate Change*, 9(1), 22–35.

Gunawardena, D. A., & de Silva, P. A. V. R. (2021). The Impact of the Fertilizer Ban on Sri Lankan Agriculture: Policy and Consequences. *Sri Lanka Journal of Agricultural Economics*, 23(1), 45–60.

Jayasuriya, S., & Jayasinghe, R. K. N. (2016). Climate Change and Drought in Sri Lanka: Impact on Paddy Yield and Agriculture. *Agricultural Systems*, 146, 1–12.

Kass, R. E., & Raftery, A. E. (1994). Bayes Factors. *Journal of the American Statistical Association*, 90(430), 773–795.

Kruschke, J. K. (2014). *Doing Bayesian Data Analysis: A Tutorial with R, JAGS, and Stan*. Academic Press.

Kumar, S., & Singh, A. (2019). Drought Impacts on Crop Production: Case Studies and Adaptation Strategies. *Journal of Environmental Management*, 241, 406–415.

Linden, A. (2015). Conducting Interrupted Time-Series Analysis for Single- and Multiple-Group Comparisons. *The Stata Journal*, 15(2), 480–500.

Lobell, D. B., & Field, C. B. (2011). Global Scale Climate-Crop Yield Relationships and the Impacts of Recent Warming. *Environmental Research Letters*, 6(1), 014010.

Marambe, B. (2022). Socio-Economic and Environmental Influences on Paddy Cultivation. *Journal of Agricultural Systems*, 18(4), 90–105.

McElreath, R. (2020). *Statistical Rethinking: A Bayesian Course with Examples in R and Stan* (2nd ed.). Chapman and Hall/CRC.

Ruppert, D., & Wand, M. P. (2009). *Semiparametric Regression*. Cambridge University Press.

Ruppert, D., Wand, M. P., & Carroll, R. J. (2003). *Semiparametric Regression*. Cambridge University Press.

Shadish, W. R., Cook, T. D., & Campbell, D. T. (2015). *Experimental and Quasi-Experimental Designs for Generalized Causal Inference*. Houghton Mifflin.

Smith, J., & Wang, L. (2020). Assessing the Combined Effects of Policy and Climate Change on Crop Yields. *Global Environmental Change*, 64, 102150.

Spiegelhalter, D. J., Best, N. G., Carlin, B. P., & van der Linde, A. (2002). Bayesian Measures of Model Complexity and Fit. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 64(4), 583–639.

Stan Development Team. (2023). *Stan Modeling Language Users Guide and Reference Manual*. Retrieved from `https://mc-stan.org`.

Thilan, A. W. L. P. (2025). Tourist Arrival Forecasting in Sri Lanka: A Bayesian Spline and Interrupted Regression Approach. *Sri Lankan Journal of Applied Statistics*, 26(1), 46–71. `https://doi.org/10.4038/sljas.v26i1.8157`.

Thilan, A. W. L. P., Peterson, E., Menéndez, P., et al. (2024). Bayesian design methods for improving the effectiveness of ecosystem monitoring. *Environmental and Ecological Statistics*, 31, 893–919. `https://doi.org/10.1007/s10651-024-00623-9`.

Thilan, A. W. L. P., Menéndez, P., & McGree, J. M. (2023). Assessing the ability of adaptive designs to capture trends in hard coral cover. *Environmetrics*, 34(3), e2802. `https://doi.org/10.1002/env.2802`.

Thilan, A. W. L. P., Fisher, R., Thompson, H., Menendez, P., Gilmour, J., & McGree, J. M. (2022). Adaptive monitoring of coral health at Scott Reef where data exhibit nonlinear and disturbed trends over time. *Ecology and Evolution*, 12(9), e9233. `https://doi.org/10.1002/ece3.9233`.

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical Bayesian Model Evaluation Using Leave-One-Out Cross-Validation and WAIC. *Statistics and Computing*, 27(5), 1413–1432.

Wagner, A. K., Soumerai, S. B., Zhang, F., & Ross-Degnan, D. (2002). Segmented Regression Analysis of Interrupted Time Series Studies in Medication Use Research. *Journal of Clinical Pharmacy and Therapeutics*, 27(4), 299–309.

Weerahewa, J., Kodithuwakku, S., & Ariyawardana, A. (2012). The Role of Rice in Sri Lankan Food Security and Rural Livelihoods. *Agricultural Economics Journal*, 43(2), 123–135.

Weerakoon, W. M. W. (2021). Policy Impacts on Fertilizer Use and Paddy Production. *Sri Lankan Agricultural Policy Review*, 8(1), 15–30.

Wheeler, T., & Von Braun, J. (2013). Climate Change Impacts on Global Food Security. *Science*, 341(6145), 508–513.

Wijeratne, M. (2018). Impact of Drought on Paddy Cultivation in Sri Lanka. *Journal of Environmental Studies*, 10(2), 56–68.

Wood, S. N. (2003). Thin Plate Regression Splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 65(1), 95–114.

Wood, S. N. (2006). *Generalized Additive Models: An Introduction with R*. CRC Press.

Wood, S. N. (2017). *Generalized Additive Models: An Introduction with R*. Chapman and Hall/CRC.

Zhao, Q., & Li, H. (2018). Interactions Between Agricultural Policies and Environmental Factors: Implications for Food Security. *Agricultural Economics*, 49(3), 367–380.