International Journal of Data Science and Analysis
2023; 9(2): 34-42
http://www.sciencepublishinggroup.com/j/ijdsa
doi: 10.11648/j.ijdsa.20230902.12
ISSN: 2575-1883 (Print); ISSN: 2575-1891 (Online)

Modeling and Forecasting the Domestic Retail Price of Teff in Ethiopia

Sisay Yohannes Gagabo

Department of Statistics, Bonga University, Bonga, Ethiopia

Email address:
sisay.john100@gmail.com

To cite this article:
Sisay Yohannes Gagabo. Modeling and Forecasting the Domestic Retail Price of Teff in Ethiopia. *International Journal of Data Science and Analysis*. Vol. 9, No. 2, 2023, pp. 34-42. doi: 10.11648/j.ijdsa.20230902.12

Received: September 13, 2023; **Accepted**: October 4, 2023; **Published**: October 28, 2023

Abstract: One of the most popular main food crops grown by the majority of Ethiopians is teff (Eragrostis teff). More than 90% of the teff consumed worldwide is grown in Ethiopia. Despite having the highest output volume, this Ethiopian cereal crop has the highest price. The major goal of this study was to estimate and predict the domestic retail price of teff in Ethiopia. The Central Statistical Agency (CSA) of Ethiopia provided the data. The average monthly domestic retail price of teff per kilogram (in birr) in Ethiopia from January 1996 to June 2023 served as the study's source of data. The data are analyzed using both descriptive and inferential statistical methods. The Statistical Packages for Social Science (SPSS Version 20.0) and R statistical tools were used to conduct the analysis. Seasonal Autoregressive Integrated Moving Average (SARIMA) model was used for modeling the average monthly domestic retail price data of teff for 27 years and forecasting for the next five years. The final model chosen, using the AIC and BIC selection criteria, was SARIMA $(2, 1, 4) \times (0, 0, 2)_{12}$, which had the minimum Mean Absolute Error (MAE) and Root Mean Square Error (RMSE). The domestic retail price of teff in Ethiopia is therefore predicted to increase relatively rapidly over the next five years, with seasonal variation. The results of this study may contribute further to the policy discussion on lowering teff prices domestically and enhancing food security. Additionally, the study is very important for managing price instability for producers, consumers, wholesalers, and national agricultural pricing policy reforms. This study also provides evidence for government policymakers on the issue of Ethiopia's exorbitant cost of living and price inflation.

Keywords: Domestic Retail Price, Teff, Time Series Data, SARIMA Model, Ethiopia

1. Introduction

Teff (Eragrostis teff), one of their favorite staple foods, is grown by the majority of Ethiopians, and the teff species originated in Ethiopia [1, 2]. Teff products' high price, on the other hand, might be used to explain their attraction. It is the most costly variety of Ethiopian grain crop and a significant source of cash for farmers. Teff has the largest revenue compared to other cereals and is even 34% bigger than coffee, in addition to giving the farmers a sizable portion of the final selling price [3].

When it comes to food intake, teff's grain is primarily used to make injera, a soft flatbread. The staple food of Ethiopia is injera, which is mostly comprised of teff. Locally known as Tikur, Sergegna, and Magna, there are primarily three different types of teff that differ in terms of quality, color, and price [4]. Since 2008, the price of teff has increased, widening the price gap with other commodities like wheat and maize. Teff's average annual price in the surplus market increased by 9% between 2015 and 2017, or 145 Birr (the Ethiopian currency equivalent of USD 5.3) per 100 kg. In January 2017, teff cost 1750 Birr ($64.1) per 100 kg, more than twice as much as wheat. As a result, teff is an economically better commodity that is primarily consumed by the wealthy [3, 5].

Teff is also an economically superior product with high income elasticity, therefore an increase in income causes an un-proportional rise in teff consumption. Some analysts argue that the recent hostilities that have taken place across the nation over the past two years, along with the Covid 19 outbreak, have had an effect on the production and distribution of teff. Within a month, a price difference of between 2500 and 3500 birr is noted. Before a month had passed; a quintal of teff cost somewhere between 5500 and

6000 birr. According to Ethiopian Media Services (EMS), which cited city residents, it is currently being sold for 9000 Birr. In some places, you may acquire it for up to 10,000 Birr [6-8].

Therefore, the primary goal of this study is to model and predict the domestic retail price of teff in Ethiopia since it may be essential for the government to ensure a program for food self-sufficiency.

2. Methodology

2.1. Data

The data for this study were the average monthly domestic retail prices of teff per kilogram (in birr) in Ethiopia from January 1996 to June 2023. These prices were determined by the Central Statistical Agency (CSA) of Ethiopia and are expressed as payments per kilogram (in birr). Following that, the entire time period was split in half. The first time frame, referred known as the "in-sample period," included the period from January 1996 to December 2020, and the data from this time frame was utilized to construct SARIMA models. The data from the "out-of-sample period," which covered the second time period was used to assess forecast accuracy, covered the months of June 2022 to June 2023.

2.2. Variables in the Study

The average domestic retail price of teff per kilogram (in birr) was the dependent variable (Y_t), while Y_{t-1}, Y_{t-2}, and so on up to Y_{t-p} = previous stationary observations (p being the number of autoregressive terms) were the predictor variables. The study used time series data.

2.3. Method of Data Analysis

2.3.1. Descriptive Statistics

The main goal of descriptive statistics is to offer clear and practical methods for summarizing interesting data aspects. In this study, descriptive statistics including mean, minimum, maximum, standard deviation, skewness, and kurtosis were used.

2.3.2. Inferential Statistics

Making inferences from sample data on the typical monthly domestic retail price of teff in Ethiopia is known as inferential statistics. As an inferential method of data analysis, time series analysis was used.

2.4. Test of Randomness and Stationary Process

2.4.1. Test of Randomness

It is important to determine whether or not the data is a random sequence. There are numerous ways to determine whether the time series data contain any systematic components or are reliant on the time t. The turning point test was employed in this scientific application because it is frequently used to determine if a batch of time-series data is truly random [9]. According to its null hypothesis, Ethiopia's average monthly domestic retail price for teff from January 1996 to June 2023 is random (there is no discernible trend).

2.4.2. Test of Stationarity

It is crucial to determine whether the study's data satisfies fundamental assumptions like stationarity before creating a Box-Jenkins modelling process. A time plot can be used to show seasonality, trends in the series' mean level or variation, long-term cycles, and other things [10].

The most common are Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) Unit Root tests; and their null hypothesis is that either the data are stationary or there is a unit root [11, 12].

2.5. Remedial Measures for Non-Stationarity

2.5.1. Transformations

It is frequently desirable or necessary to modify a time series data set before employing sophisticated techniques. Differentiation, power transformation, dividing seasonal standard deviation, and subtracting seasonal averages are a few of the various transformation procedures [13]. Commonly utilized differencing and Box-Cox power transformation methodologies were used in this investigation.

2.5.2. Stationary Through Differencing

The first difference of a time series is the series of changes from one period to the next. If Y_t denotes the value of the time series Y at period t, then the first difference of Y at period t is equal to Y_t-Y_{t-1}.

$$\Delta Y_t = Y_t - Y_{t-1} \tag{1}$$

$$\Delta^2 Y_t = Y_t - Y_{t-1} - Y_{t-2} \tag{2}$$

Where, ΔY_t and $\Delta^2 Y_t$ are the first and second difference of Y_t respectively [14].

2.6. SARIMA Model

To find the best model for predicting the price data that corresponds to the real domestic retail price of teff in Ethiopia, Seasonal Autoregressive Integrated Moving Average (SARIMA) model was tested. The identification, parameter estimates, diagnostic evaluation, and forecasting phases of the Box-Jenkins Seasonal ARIMA model fitting are the key steps.

The SARIMA model is preferred when any seasonal behavior in the series is suspected because the ARIMA model is ineffective for those series with both seasonal and non-seasonal behavior, such as incorrect order selection. The notation (p, d, q) × (P, D, Q)$_S$ is used to represent the multiplicative SARIMA model [15].

The corresponding lag form of the model is:

$$\emptyset(L)\varphi\left(L^S\right)\left(1-L\right)^d\left(1-L^S\right)^D Y_t = \theta(L)\vartheta\left(L^S\right)\varepsilon_t \tag{3}$$

This model includes the following Auto Regressive (AR) and Moving Average (MA) characteristic polynomials in L of order p and q respectively:

$$\emptyset(L)=1-\emptyset_1 L-\emptyset_2 L^2-\ldots-\emptyset_{p-1} L^{p-1}-\emptyset_p L^p \quad (4)$$

$$\theta(L)=1-\theta_1 L-\theta_2 L^2-\ldots-\theta_{q-1} L^{q-1}-\theta_q L^q \quad (5)$$

Also Seasonal polynomial functions of order P and Q respectively as represented below:

$$\varphi(L^S) = 1 - \varphi_1 L^S - \varphi_2 L^{2S} - \cdots - \varphi_{P-1} L^{(P-1)S} - \varphi_P L^{PS} \quad (6)$$

$$\vartheta(L^S) = 1 - \vartheta_1 L^S - \vartheta_2 L^{2S} - \cdots - \vartheta_{Q-1} L^{(Q-1)S} - \vartheta_Q L^{QS} \quad (7)$$

Where: $\{Y_t\}$ - the observable time series, $\{\varepsilon_t\}$ - white noise series, (p, d, q) - order of non-seasonal AR, differencing and non-seasonal MA respectively, (P, D, Q) - order of seasonal AR, differencing and seasonal MA respectively, L - lag operator $L^k Y_t=Y_{t-k}$, and S - seasonal order for example S=12 for monthly data.

2.7. Modelling Approaches

The first stage in creating the SARIMA model is to determine whether the series meet the stationarity requirement, which calls for a time-invariant mean, variance, and co-variance.

It is possible to ascertain the orders p, q, P, and Q by examining the patterns of the Auto Correlation Function (ACF) and Partial Autocorrelation Function (PACF). These employ data on internal correlation between observations made for a time series at various intervals to give a sense of seasonal and non-seasonal lags. At the non-seasonal and seasonal levels, the ACF and PACF, respectively, both have spikes and cut off at lag k and lag ks. The quantity of statistically significant spikes determines the model's ranking [16, 17].

The autocorrelation coefficient (r_k) with the order of k generally cuts off for an MA (0, d, q) process and is not significantly different from zero after lag q. An AR term may be required to model the time series if r_k tails off but does not terminate. The autocorrelation coefficient (r_k) truncates and is not significantly different from zero after lag q + sQ when the process is SARIMA (0, d, q) * (0, D, Q). If r_k begins to weaken at lags that are multiples of s, a seasonal AR component is likely present. The PACF (ϕ_{kk}) with the order of k truncates for an AR (p, d, 0) process and is not appreciably different from zero after lag p. If ϕ_{kk} tends to zero gradually, it is implied that an MA term is necessary. When a SARIMA (p, d, 0)*(P, D, 0) process is involved, ϕ_{kk} cuts off and is not appreciably different from zero after lag p + sP. A seasonal MA component should be included in the model if ϕ_{kk} dampens out at lags that are multiples of s [18].

The ACF and PACF may provide a number of distinct models, each of whose parameters is estimated by the maximum likelihood approach. The optimal model is determined using the Box-Pierce-Ljung test, the "auto.arima" function in R, and the model with the lowest AIC and BIC selection criterion values. Residual diagnostic testing is the final phase in the model selection process, and if the model passes these diagnostic checks, it can be utilized to forecast.

Table 1. *Behaviour of ACF and PACF for Non-seasonal ARMA (p, q).*

		AR(p)	**MA(q)**	**ARMA(p, q)**
Non-seasonal ARMA (p, q)	ACF	tails off at lag k, k=1, 2, 3,...	cuts off after lag q	tails off
	PACF	cuts off after lag p	tails off at lags k k=1, 2, 3,...	tails off

Table 2. *Behaviour of ACF and PACF for seasonal ARMA (P, Q).*

		AR(P)s	**MA(Q)s**	**ARMA(P,Q)s**
Pure-seasonal ARMA (P, Q)	ACF	tails off at lag ks, k=1, 2, 3,...	cuts off after lag Qs	tails off at ks
	PACF	cuts off after lag Ps	tails off at lags ks k=1, 2, 3,...	tails off at ks

Note that: In this context "tails off" means "tend to zero gradually" and "cut off" means "disappear or is zero".

2.8. Assessment of Forecast Accuracy

This study included measures of forecast accuracy that are often used, such as Mean Absolute Error (MAE) and Root Mean Square Error (RMSE), and it was determined that the best model was the one with the lowest MAE or RMSE [19-21].

$$\text{MAE} = \frac{1}{n} \quad (8)$$

$$\text{RMSE} = \frac{1}{n}\sum_{t=1}^{n} (\varepsilon_t)^2 \quad (9)$$

Where the difference between the actual observation (the average domestic retail price of teff per kilogram (in birr)) Y_t and the forecasted value $\hat{y}_t$ is the forecast error $\varepsilon_t=Y_t-\hat{y}_t$, and n is the sample size (the length of the time series).

3. Results

3.1. Descriptive Statistics

A total of 330 observations on the monthly domestic retail price of teff in Ethiopia from January 1996 to June 2023 were used as the study's data. The descriptive statistics of the average monthly domestic retail price of teff per kg (in birr) are shown in Table 3.

According to the empirical finding in Table 3, the domestic teff price per kg was 14.02 on average per month, with a standard deviation of 14.76. Furthermore, the mean and standard deviation values are roughly similar. The exponential distribution is one instance when the mean and standard deviation are equal [22]. The multiplicative approach is acceptable if the time series has an exponential distribution over time [23]. Therefore, this data scenario is a suitable starting point for quickly choosing an acceptable model.

According to published research, the agricultural price distribution demonstrates the following characteristics: skewness and leptokurtosis [24]. Because the monthly price of the teff series often has asymmetric distributions skewed to the right, as indicated by the coefficient of skewness of 2.10, the evidence for price series in Table 3 suggests positive skewness and longer tails than does the normal distribution for the series. Additionally, the monthly domestic price of the Teff series has a leptokurtic distribution, as indicated by the excess kurtosis coefficient of 5.87. Teff's average domestic retail price per kilogram in Ethiopia was therefore not distributed normally.

Table 3. *Summary of the average monthly domestic retail price of teff per kg (in birr).*

Statistic	Minimum	Maximum	Mean	Standard deviation	Skewness	Kurtosis
Price	1.79	100.00	14.02	14.76	2.10	5.87

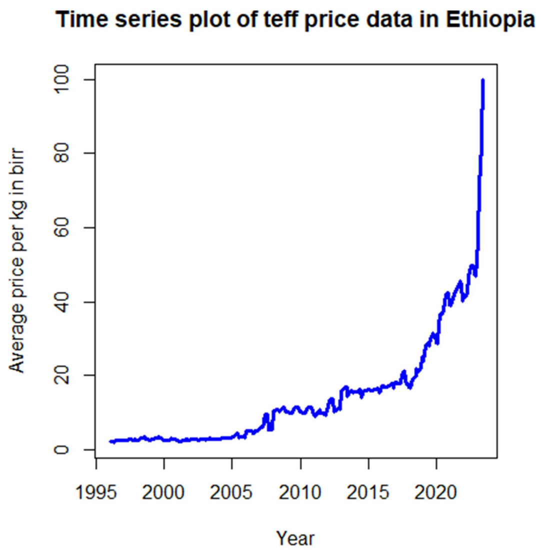

Figure 1. *The domestic retail price trend of teff from January 1996 to June 2023.*

In Figure 1, the Y-axis represents the average monthly domestic price per kg of teff, and the X-axis represents the years in the sample period, which runs from January 1996 to June 2023. The monthly domestic price of teff per kg under research exhibits an increasing tendency across the study period from January 1996 to June 2023, as can be seen from the figure above. After the year 2020, a particularly sharp increase in the domestic price of teff was noted in the nation. The image also demonstrates that the pricing trend is increasing exponentially over time, which is consistent with the findings in Table 3 above.

Because of a noticeable increase trend in the average monthly domestic retail price of teff in Ethiopia, it is clear from Figure 1 that the data are non-stationary. To create a stationary series, the mean and variance should be changed to make the numbers vary more or less over time. In order to achieve stationarity, the mean and variance may therefore need to be adjusted using differencing and transformation, respectively [13, 14]. But to make sure of that, we perform the Augmented Dickey-Fuller and Phillips-Perron (PP) Unit Root tests.

3.2. Test of Randomness

This test is intended to determine whether the time series data have any systematic components or whether they are dependent on the passage of time t. The assumption that the data are independent and uniformly distributed is the main source of worry [25]. There are four widely used tests of randomization for this purpose. The rank test, turning point test, difference single tests, and phase length test are among them. The researcher employed the most popular turning point test in this occasion [9]. Additionally, its null hypothesis states that teff's domestic average retail price fluctuates at random.

Table 4. *Turning point test for randomness of price of teff.*

Time series data	Randomness test	Critical value at 5% level of significance	P-value
Price of teff	Turning point test	-8.694	0.000

Given that the p-value is 0.000 in Table 4, we reject the null hypothesis. The data on the average monthly price of teff is therefore not random at the 5% level of significance. This demonstrates a systematic pattern in the data and also suggests that the data's distribution is not stationary. Accordingly, the data must be appropriately transformed into a stationary form.

3.3. Unit Root Test for Stationarity

A series' behavior and features, such as the persistence of shocks and erroneous regressions, can be significantly impacted by its non-stationarity. In other words, even if two variables are completely unrelated, a regression of one on the other could have a high R^2 if the variables are trending over time [12]. In order to determine whether the study's data were stationary, the Augmented Dickey-Fuller (ADF) and Phillips-Perron (PP) unit root tests were run. The average domestic retail price of teff in Ethiopia from January 1996 to June 2023 has a unit root, according to the null hypothesis of this test.

Table 5. *ADF and PP test for the average monthly domestic retail price of teff.*

Time series data	Stationarity test	Critical value at 5% level of significance	P-value
Price of teff	Augmented Dickey-Fuller (ADF) test	4.222	0.99
	Phillips-Perron (PP) Unit Root test	25.637	0.99

The price of the teff under examination did not appear to be stationary, according to the findings of the stationarity test in Table 5. This is due to the fact that the relevant p-values from the ADF and PP tests were more than 5% level of significance, failing to reject the null hypothesis of non-stationarity. As a result, there is no data to support the non-stationarity null hypothesis at the 5% level of significance. After remedial actions were taken to make the series stationary, which is necessary for further analysis, the average monthly domestic retail price of teff, however, looked stationary.

3.4. Remedial Measures for Non-Stationarity

A crucial premise of time-series models is stationary. Because the series performs similarly throughout a shorter time span when it is stationary, its statistical characteristics do not change with time. This is referred to as homogeneity. When the observations are converted via a Box-Cox transformation using a lambda value of 0.5, they are frequently made to satisfy these conditions even if the constant variance and normality assumptions are not satisfied [26].

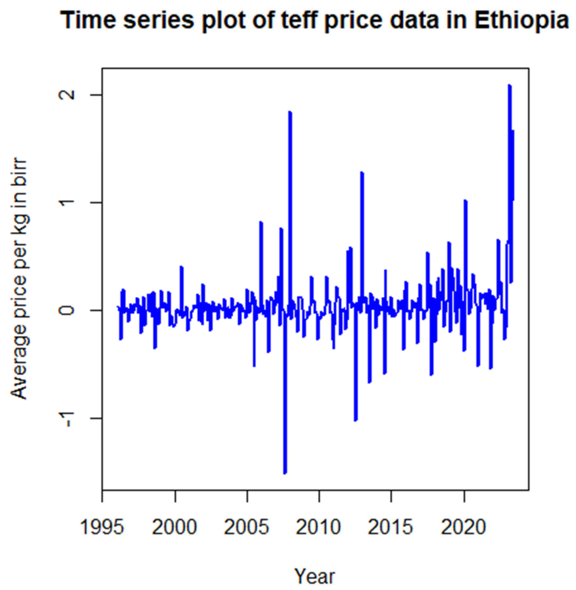

Figure 2. *Stationary time series of teff price after remedial measures.*

The transformed data becomes a stationary series after being differentiated with a lambda value of 0.5 in lag 1, as shown in Figure 2. This plot demonstrates how the series deviates from the mean. This shows that it stays close to the mean and that the time plots look to be consistent throughout. The time series is stationary as a result, indicating that there is little separation between changes in the average monthly domestic retail price of teff.

3.5. SARIMA Modelling Approaches

For the monthly series, information regarding the seasonal and non-seasonal AR and MA operators should be gathered using autocorrelation functions (ACF) and partial autocorrelation functions (PACF) [18].

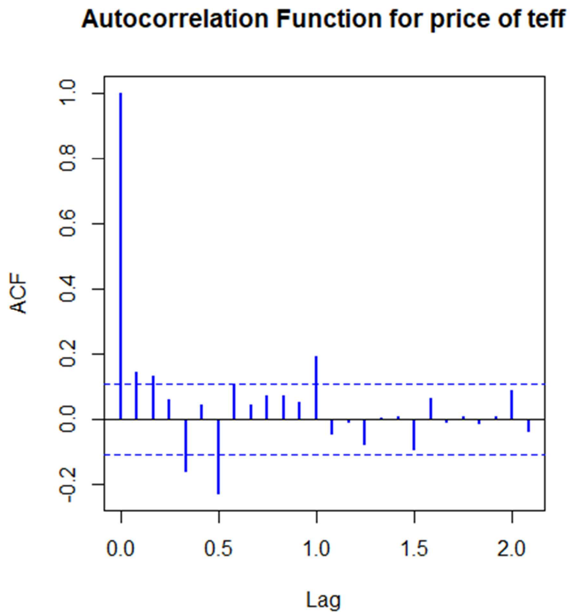

Figure 3. *ACF on transformed data of domestic retail price of teff.*

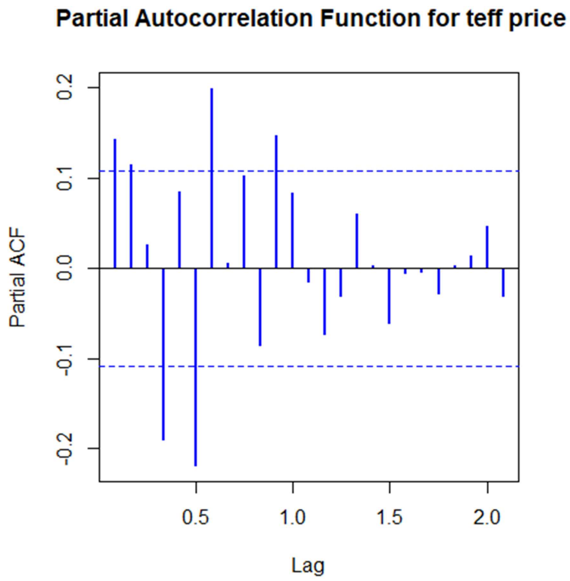

Figure 4. *PACF on transformed data of domestic retail price of teff.*

ACF and PACF graphs for the data reflect the periodicity of the data and may suggest the necessity for non-seasonal and seasonal components in the model, according to visual inspections of the aforementioned Figures 3 and 4. As a result, it is challenging to determine the order of a mixed model in this circumstance. The auto.arima function in the statistical software package R is the finest tool for determining the order of a mixed model [27]. It selects the best model form different candidate models with the lowest AIC and BIC selection criterion values. The final model chosen was SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$.

Table 6. *Parameter estimates of the final model.*

Type	Estimate	Standard error	Z-value	P-value
AR 1	-1.364399	0.129257	-10.5557	0.00000
AR 2	-0.642676	0.114852	-5.5957	0.00000
MA 1	0.601170	0.131165	4.5833	0.00000
MA 2	-0.522716	0.080349	-6.5055	0.00000
MA 3	-0.627096	0.124846	-5.0230	0.00000
MA 4	-0.336641	0.068493	-4.9150	0.00000
SMA 1	0.218543	0.068033	3.2123	0.001317
SMA 2	0.100693	0.061790	1.6296	0.103185

The empirical findings of parameter estimations for the SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ model are shown in Table 6. The parameters of the final model differ significantly from zero, with the exception of the final value for the seasonal component. This led to the conclusion that, for predicting the price of teff over time, the SARIMA model is preferable to the ARIMA model.

3.6. Final Model Diagnostics

The fitted model residuals were evaluated in order to assess the goodness of fit of the fitted SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ model in the modeling of the average monthly domestic retail price of teff data. The model residual studies were aided by the use of the standardized residuals, ACF of residuals, and Ljung Box p-values [28].

Table 7 provided an overview of the Ljung Statistic to check for auto-correlations in the residuals. There were no auto-correlations in the residuals, as indicated by the Ljung Statistic p-value being greater than 0.5. Since the residuals could not be distinguished from a white noise series, the SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ was thought to have a good fit for the average monthly domestic retail price of teff data.

Table 7. *The Ljung Statistic of final model.*

Final model SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$	Model Adequacy test	Critical value at 5% level of significance	P-value
Residuals	Box-Pierce-Ljung test	0.0011952	0.9724

Shapiro-Wilk normality test's null hypothesis was that the residuals from the SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ model are not normally distributed [29]. The residuals from the fitted model were subjected to a normality test in Table 8, which indicated that the residuals were normally distributed. The SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ model seems to fit the average monthly domestic retail price of teff data more closely, according to the normality distribution of the fitted model residuals.

Table 8. *Normality test of residuals of final model.*

Final model SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$	Normality test	Critical value at 5% level of significance	P-value
Residuals	Shapiro-Wilk Normality test	0.83875	0.0000

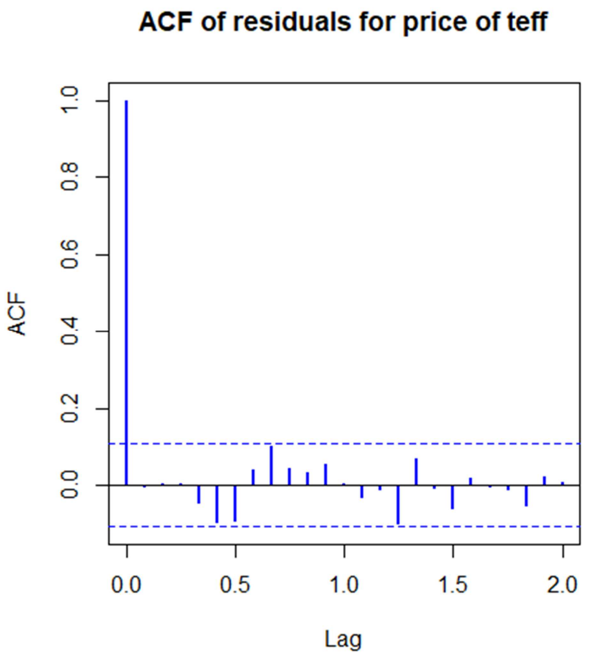

Figure 5. *ACF of residuals for final model of teff price.*

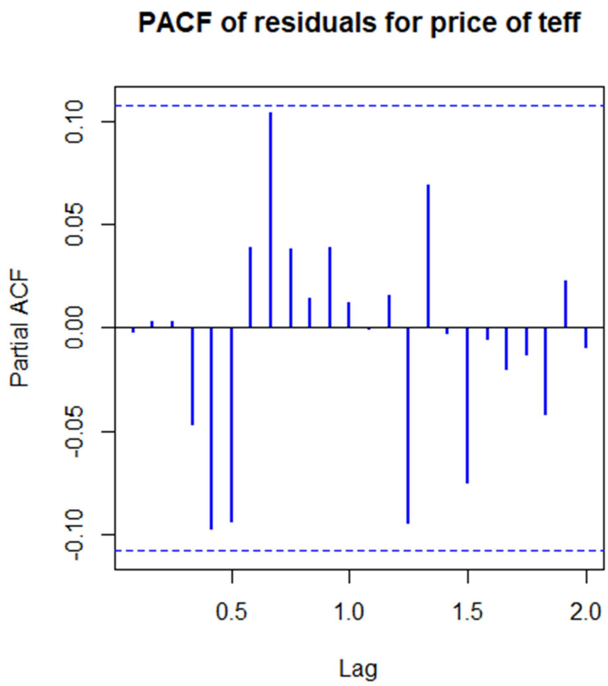

Figure 6. *PACF of residuals for final model of teff price.*

The diagnostic plots for the residual ACF and PACF of the fitted SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ model were shown in Figures 5 and 6. The goal of the model verification was to evaluate if the residuals of the fitted model had any systematic patterns that might still be eliminated to enhance the selected SARIMA. This was accomplished by looking at the residuals' autocorrelation and partial auto-correlation.

The plots revealed no evidence of serial correlation in the residuals of the series, which had a mean close to zero and fell below the tolerance line in the p-value plot. The fitted model was appropriate and good in the modeling of the average monthly domestic retail price of teff data because the ACF and PACF of the residuals do not exceed a significant bound from lag 1 to the end.

Therefore, all of the results suggested that the fitted model successfully captured the data (it accounted for all the information provided) and could be used to forecast the domestic retail price of teff in the future.

3.7. Forecast of Price of Teff in Ethiopia

To predict the domestic retail price of teff over time, fitted ARIMA (2, 1, 4) and SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ were employed.

Table 9. *Evaluation of forecasting for the final model.*

Model	Measures of forecast accuracy	Value
ARIMA (2, 1, 4)	Mean Absolute Error (MAE)	60.66594
	Root Mean Square Error (RMSE)	62.96404
SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$	Mean Absolute Error (MAE)	60.60721
	Root Mean Square Error (RMSE)	62.87839

The average domestic retail price of teff data predicted for the fitted model was visualized in Figure 7. The model forecasting was done using the in-sample range data [30]. The straight line represents the model forecast, which was at the start of the forecast range. Additionally included are the confidence intervals for both the model forecast and the in-sample range data.

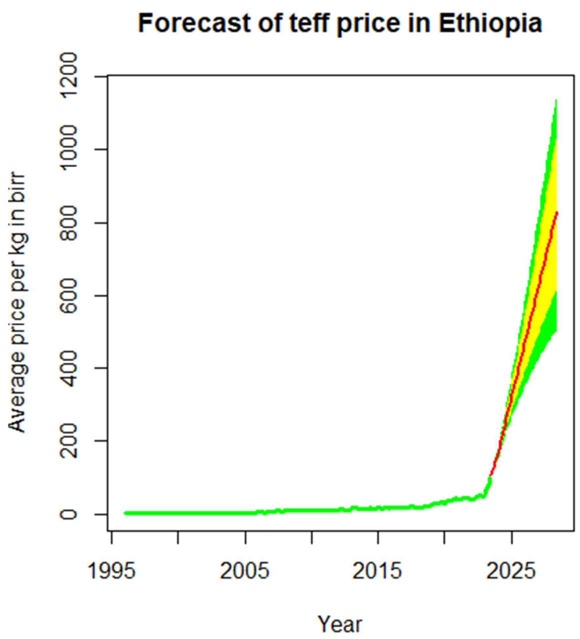

Figure 7. *Forecast for price of teff in Ethiopia for the next five years.*

Finally, we compare the predicted values of ARIMA (2, 1, 4) and SARIMA (2, 1, 4) × $(0, 0, 2)_{12}$ with the observed values to determine which of the two fitted models is more sparse. Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) were employed to assess their performance. In Table 9, the ARIMA model's MAE and RMSE were 60.66594 and 62.96404, respectively, whereas the SARIMA model's MAE and RMSE were 60.60721 and 62.87839. A conclusion that the SARIMA model is much superior than the ARIMA model in the modeling of domestic retail price of teff data can be drawn from the fact that the MAE and RMSE values of the SARIMA model were smaller than the MAE and RMSE values of the ARIMA model.

4. Conclusion

The average monthly domestic retail price of teff per kilogram (in birr) in Ethiopia from January 1996 to June 2023 served as the study's source of data. These prices are expressed as payments per kilogram (in birr) and were established by the Central Statistical Agency (CSA) of Ethiopia. Teff's typical domestic retail price is steadily rising in Ethiopia. As opposed to earlier years, the statistics for 2020–2023 show a high degree of variation in the domestic retail price of teff, suggesting that the cost of teff varies significantly from month to month.

The following aspects are outlined in light of the analysis done in this study on the average domestic retail price of teff: The average domestic retail price of teff is time-dependent, according to the time series plot. The data has been converted into a stationary time series through the first difference of the transformed data before applying the series models. The future price of teff in Ethiopia was predicted using the two ARIMA and SARIMA models, which were compared. The SARIMA (2, 1, 4) (0, 0, 2) model was the one that performed the best in predicting the price of teff. It had the lowest mean absolute error (MAE) and root mean square error (RMSE), and was picked using the AIC and BIC. With the exception of the final value for the seasonal component, the final model's parameters differ significantly from zero. This led to the conclusion that the SARIMA model is superior to the ARIMA model for predicting the price of teff over time.

The most accurate model has been applied to forecast the price of teff in Ethiopia in the future. Therefore, it is predicted that over the next five years, the domestic retail price of teff in Ethiopia will rise very quickly. The rapid rise in domestic demand brought on by rising incomes, population, and urbanization is responsible for the high price. To improve agronomic results, the Ethiopian government must ensure ongoing public investment in teff research.

Availability of Data and Materials

The author has accessed the data from the Ethiopian Statistical Agency (CSA). The dataset supporting the conclusions of this article is available by contacting the author.

Competing Interests

The author declares that he has no competing interests.

Author Contribution

The author has designed the study, analyzed the data, drafted the manuscript, and critically reviewed the article.

Acknowledgments

The author wishes to express his sincere appreciation to the Central Statistical Agency of Ethiopia for making the data public and for providing the monthly domestic retail price data.

References

[1] R. Dibaba Wake, A. Hagos Mesfin, C. Yirga, and E. Habte, "Determinants of Improved Teff Varieties Adoption and Its Impact on Productivity: The Case of Non-Traditional Teff Growing Areas of Western Ethiopia," *J. Nat. Sci. Res. www.iiste.org ISSN*, vol. 8, no. 22, 2018, [Online]. Available: www.iiste.org

[2] T. Maychew, "The impact of row planting of teff crop on rural Household income: A case of Mekelle University College Business and Economics Department of Economics The impact of row planting of teff crop on rural Household income: A case of Tahtay Maychew wereda, T," vol. 19, no. June, 2020.

[3] E. Tadele and T. Hibistu, "Empirical review on the use dynamics and economics of teff in Ethiopia," *Agric. Food Secur.*, vol. 10, no. 1, pp. 1–13, 2021, doi: 10.1186/s40066-021-00329-2.

[4] B. Gizaw, "Traditional Knowledge on Teff (Eragrostistef) Farming Practice and Role of Crop Rotation to Enrich Plant Growth Promoting Microbes for Soil Fertility in East Showa: Ethiopia," *Agric. Res. Technol. Open Access J.*, vol. 16, no. 5, 2018, doi: 10.19080/artoaj.2018.16.556001.

[5] B. Minten, T. Seneshaw, E. Ermias, and K. Tadesse, "Ethiopia's Value Chains on the Move: The Case of Teff. ESSP (Ethiopian Strategic Support Program)," no. April, 2013.

[6] B. Minten, A. S. Taffesse, and P. Brown, *The economics of teff: Exploring Ethiopia's biggest cash crop*. Intl Food Policy Res Inst, 2018.

[7] S. A. Omer and N. A. Hassen, "Impacts COVID-19 pandemic diseases on Ethiopian agriculture, food systems, industries, and mitigation and adaptation strategy," *Electron. J. Educ. Soc. Econ. Technol.*, vol. 1, no. 1, pp. 18–33, 2020.

[8] E. Teka, "The price of Injera in several Ethiopian cities goes up | Addis Zeybe - Digital Newspaper," 2021. https://addiszeybe.com/featured/addis-ababa/market/economy/the-price-of-injera-in-several-ethiopian-cities-goes-up (accessed Sep. 18, 2023).

[9] R. H. Shumway, D. S. Stoffer, and D. S. Stoffer, *Time series analysis and its applications*, vol. 3. Springer, 2000.

[10] C. Chatfield and D. L. Prothero, "Box-Jenkins seasonal forecasting: Problems in a case-study," *J. R. Stat. Soc. Ser. A*, vol. 136, no. 3, pp. 295–315, 1973.

[11] D. A. Dickey and W. A. Fuller, "Likelihood ratio statistics for autoregressive time series with a unit root," *Econom. J. Econom. Soc.*, pp. 1057–1072, 1981.

[12] P. C. B. Phillips and P. Perron, "Testing for a unit root in time series regression," *Biometrika*, vol. 75, no. 2, pp. 335–346, 1988.

[13] A. Nielsen, *Practical time series analysis: Prediction with statistics and machine learning*. O'Reilly Media, 2019.

[14] D. Kwiatkowski, P. C. B. Phillips, P. Schmidt, and Y. Shin, "Testing the null hypothesis of stationarity against the alternative of a unit root: How sure are we that economic time series have a unit root?," *J. Econom.*, vol. 54, no. 1–3, pp. 159–178, 1992.

[15] J. Brownlee, "A gentle introduction to sarima for time series forecasting in python," *Mach. Learn. Mastery*, 2018.

[16] E. AIDOO, "Forecast performance between sarima and setar models: An application to ghana inflation rate." 2011.

[17] R. H. Shumway and D. S. Stoffer, "Time series regression and exploratory data analysis," *Time Ser. Anal. its Appl. With R examples*, pp. 48–83, 2006.

[18] W. Vandaele, "Applied time series and Box-Jenkins models," Introduction to time series, pp. 67–123, 1988.

[19] N. Vandeput, "Forecast KPI: How to Assess the Accuracy of a Product Portfolio Forecasting KPIs such as MAPE, MAE, and RMSE are not suited to assess the accuracy of a product portfolio. Let's take a look at a few new metrics: MASE, RMSSE, WMASE, and WRMSSE.".

[20] T. Chai and R. R. Draxler, "Root mean square error (RMSE) or mean absolute error (MAE)," *Geosci. Model Dev. Discuss.*, vol. 7, no. 1, pp. 1525–1534, 2014.

[21] R. J. Hyndman and G. Athanasopoulos, *Forecasting: principles and practice*. OTexts, 2018.

[22] M. Hardy, "Mean and std deviation of a population equal?," 2023. https://math.stackexchange.com/questions

[23] A. Omidifar, "Transformers in Time Series," 2019. https://www.linkedin.com/pulse/considerations-time-series-part-i-andrea-schmidt (accessed Sep. 18, 2023).

[24] B. B. Anjullo, "Modeling Domestic Price Volatility for Cereal Crops in Ethiopia," *Int. J. Data Sci. Anal.*, vol. Vol. 7, no. No. 6, pp. 139–149, 2021, doi: 10.11648/j.ijdsa.20210706.12.

[25] N. Beck and J. N. Katz, “Modeling dynamics in time-series–cross-section political economy data,” *Annu. Rev. Polit. Sci.*, vol. 14, pp. 331–352, 2011.

[26] K. Yürekli and A. Kurunç, “Simulation of drought periods using stochastic models,” *Turkish J. Eng. Environ. Sci.*, vol. 28, no. 3, pp. 181–190, 2004.

[27] R. J. Hyndman and Y. Khandakar, “Automatic time series forecasting: the forecast package for R,” *J. Stat. Softw.*, vol. 27, pp. 1–22, 2008.

[28] G. E. P. Box and D. A. Pierce, “Distribution of residual autocorrelations in autoregressive-integrated moving average time series models,” *J. Am. Stat. Assoc.*, vol. 65, no. 332, pp. 1509–1526, 1970.

[29] G. S. Mudholkar, D. K. Srivastava, and C. Thomas Lin, “Some p-variate adaptations of the Shapiro-Wilk test of normality,” *Commun. Stat. Methods*, vol. 24, no. 4, pp. 953–985, 1995.

[30] N. Amjady and F. Keynia, “Day ahead price forecasting of electricity markets by a mixed data model and hybrid forecast method,” *Int. J. Electr. Power Energy Syst.*, vol. 30, no. 9, pp. 533–546, 2008.