P-ISSN 2664-0422	Pakistan Social Sciences Review	October-Dec 2024, Vol. 8, No. 4
O-ISSN 2664-0430	https://doi.org/10.35484/pssr.2024(8-IV)61	[672-684]

Pakistan Social Sciences Review
www.pssr.org.pk

RESEARCH PAPER

Time Series Analysis of Soybeans Prices in Pakistan Applying Symmetric and Asymmetric GARCH Models with Normal and Non-Normal Innovations

[1]Tahira Bano Qasim * [2]Sofia Akbar and [3]Maryam Siddiqa

1. Associate Professor, Department of Statistics, The Women University Multan, Punjab, Pakistan
2. M. Phil Scholar, Department of Statistics, The Women University Multan, Punjab, Pakistan
3. Lecturer, Department of Statistics, The Women University Multan, Punjab, Pakistan

***Corresponding Author:** tahirabanoqasim@yahoo.com

ABSTRACT

Forecasting and volatility modeling are important tools for all agricultural and financ sectors. The core aim of this study is to compare the performance of symmetric an asymmetric Generalized Autoregressive Conditional Heteroscedasticity (GARC models under both normal and non-normal error distributions to forecast soybean dai prices in Pakistan. The ARIMA models are applied as mean models along with fo GARCH models (GARCH, EGARCH, TGARCH and PARCH) with three err distributions (Normal, Student-t and GED). The ACF and PACF of residual and squar residuals are used as diagnostics to check the appropriateness of the models. Based on t empirical findings it is concluded that improvement in the overall estimation is achiev using asymmetric GARCH models as the conditional variance. Moreover, TGARC model with student-t distribution outperforms the other models in forecasting soybe prices in Pakistan. These results provide valuable insights for stakeholders an policymakers in choosing a suitable forecasting model for soybean prices in Pakistan. T government should take steps to develop high-yielding latest production technology an strategies to increase the production of soybeans.

KEYWORDS ARMA, Asymmetric, EGARCH, GARCH, GED, Normal Distribution, PARCH, Soybean Prices, Student-t Distribution, Symmetric, TAGRCH

Introduction

Soybean is one of the major crops and the primary source of food for the vast population of the world. With a continuous increase in global yearly production, soybean has become an increasingly important agricultural product. Soybeans as one of the world's most important and quickest crops, contribute significantly to the improvement of human diets in levels of both protein and energy consumption, the crops have been well-positioned to meet the fast-growing demand for vegetable oil, and animal feeding in developing countries. Soybean is also known as the "golden bean" because, when compared to other oilseed crops, it has a high nutritive value, excellent flavor, and a large amount of oil and proteins. The best source of soybean oil provides important nutrients for human development and growth. Soy is also used in animal feeding, biodiesel, soy crayons, bio-composite, carpets, candles, concrete, soy foams, soy ink, cleaners, and soy tries. Soybean is worldwide produced as the fourth main crop. Moreover, recent national and international agricultural policies appear to limit smallholder soybean production in low-income, nutritious nations while supporting worldwide soybean and related product trading.

Pakistan is an agricultural country, and agriculture plays a major role in the country's development and economy. Agriculture is the backbone of Pakistan's economy, and it affects around 70% of the population directly or indirectly. Soybeans are short-season crop that takes 90 to four months to mature, depending on seed variety and weather conditions. This crop requires less fertilization than cash crops such as rice, sugarcane, maize, and many others. Soybean is produced on a total area of 120.48 thousand acres worldwide but these are grown on a smaller scale in Pakistan (see Dilawaiz, 2010). Therefore, the Soybean oilseed crop has a gap in domestic production and supply. However, Pakistan spends a lot of its foreign exchange on the trade-in soybeans oil and oilseed-based nutrition to specialize in its production. During FY2018-2019 (July -March), 1.445 billion US dollars were spent on oil products to supply 80% of the total national annual domestic purposes, whereas local products produced only 20% of overall growth (see GOP 2019). Even though Pakistan is a challenging country to export through because of significant taxes and governmental and religious and transportation advantages, with competition from countries. In Pakistan, soybean imports have increased significantly, then per capita consumption remains very low, indicating that more imported products are on the way among 2014-2015 and 2018-2019, imports have grown by 455% and 478% including both, with imports predicted to rise, as according to USDA (see Barrett and LLC, 19 Feb, 2020). Soybean price forecasting assists farmers and industries in planning future farming activities and budgeting, which is heavily reliant on anticipated future pricing. As a result, estimating future soybean prices has become an important part of price policy. Price estimations, on the other hand, are crucial in making better decisions and will play a key role in coordinating global soybean supply and demand.

The basic purpose of this study is to analyze which model of time series gives us better forecasts for soybean daily prices in Pakistan. For achieving our target, the ARIMA modeling technique is applied to mean modeling. The variations and asymmetries exhibited by soybean daily price series are accommodated by applying GARCH models with normal and non-normal innovations. In the context of non-normal distributions, students-t and generalized error distributions are utilized. The main motivation of this study is to model the soybean price series in Pakistan to capture the volatility clustering, asymmetric response to good and bad news and fat tails of the data which was not done earlier as depicted by literature.

Literature Review

Different time series models are applied in the analysis of soybean prices across different areas: Tinker at el. (1989) applied a univariate time series model ARIMA and multivariate VAR models to forecast the prices of the soybean complex (soybeans, soybean meal, and soybean oil). Lence and Hayes (2002) applied GARCH-type models in forecasting soybean price volatility and found that EGARCH models provide better forecasts as compared to simple GARCH model. Felipe et al. (2012) used the ARIMA modeling approach to evaluate the daily soybean pricing series in the North of Parana and explained their behavior in short-term period projections. Their empirical findings revealed that ARIMA (5, 0, 0) is the best model for predicting soybean prices among the models evaluated. Arı and Tüfekci (2014) applied GARCH, EGARCH, and GJR-GARCH with different distributional assumptions in forecasting soybean prices. Their findings are consistent with the results that asymmetric GARCH models outperformed symmetric one when using non-normal distributions for error terms. Risso et al. (2011) analyzed soybean prices and suggested a simple theoretical model for determining the equilibrium price. They observed a co-integration relationship between monthly U.S. soybean prices, the Chinese Yuan per Dollar exchange rate, and Chinese real GDP. Felipe et al. (2012) used the ARIMA modeling approach to evaluate the daily soybean pricing series in the North of Parana and explained their behavior in short-term period projections. Wiles and Enke (2015) explored the moving average convergence divergence parameter values from conventionally used integers to values that rely on the soybean prices at the entry and exit points Abraham at el. (2020) studied the Artificial Neural Networks (ANN) and evaluate classical methods of Time Series Analysis to forecast Brazilian production and soybean harvest area, and yield. Xu et al. (2021) represented the research object of the Dalian commodity futures market in china in soybean futures and studied of correlation between commodities prices attributes, trading volume, and open position. Chi (2021) applied the hybrid model of the SARIMA and NARNN models in predicting process data of monthly global price of soybeans for the period from January 1990 to January 2021. The findings revealed that the Hybrid-LM model performed better than the NARNN-LM model and the SARIMA model by means of MSE.

In the context of methodology that is related to the study are: Peters (2001) analyzed and compared the forecasting of the two main European stock indexes (FTSE 100 and DAX 30)by using four symmetric GARCH models namely TGARCH, EGARCH, GJR, and PARCH and three distributions (Normal, Student-t, and Skewed Student-t). McMillan and Speight (2004) compared the performance of GARCH models with normal, Student's t, and GED distributions in forecasting stock market volatility and found that non-normal distributions consistently outperformed the distribution in improving forecast accuracy. Shamiri at el. (2007) presented the study of modeling and forecasting the two Asian stock indices (KLCI and STI) by using daily data over 14 years. They estimated three models (GARCH, EGARCH, and GJR-GARCH), with three distributions Normal, Student -t distribution, and Generalized Error distribution (GED). Pasha et al.(2007) compared the GARCH family of models with three distributions(Normal, student's t and GED) to forecast KSE100 index in Pakistan. Their results support the use of asymmetric GARCH models with non-normal distributions. Sharma et al. (2021) compared the linear and non-linear GARCH models in forecasting the volatility of stock markets of Indonesia, China, India, Brazil and Mexico. Their findings support the use of linear GARCH (1, 1) model over non-linear GARCH models for forecasting volatility indicating the insignificance of effect of leverage. In the present study we have compared the GARCH models with three error distributions.

Material and Methods

This research is about the analysis of daily soybean prices in Pakistan. For the period from 1995 to 2022, time series data for soybean prices in Pakistan in U.S. dollars per bushel was collected from the website "www.macrotrend.net." The whole data consists of 7074 observations, of which 5728 observations are used to estimate the model and the remaining 1306 observations are used to evaluate forecasting. Statistical software Excel and EViews 9 are used for analysis purposes.

The study's methodology is summarized as follows steps:

- Data visualization
- ACF, PACF, and the Dickey and Fuller Unit test are used to check for stationarity.
- Applying ARIMA models for estimating the mean model.
- Estimating the structure of conditional heteroscedasticity underlying the series under consideration.
- Evaluating the forecasts obtained by applying ARMA models with symmetric and asymmetric GARCH models with normal and non-normal distributional assumptions.

- The diagnostics of the residuals and squared residuals are checked by the correlogram of the ACF, PACF.
- Model selection criterion is based on Akaike info criterion (AIC), Schwarz criterion (SC), Hannan-Quinn criterion (HQ) and Maximum log-likelihood
- Evaluating the forecast using the mean absolute error, mean absolute percentage error, and RMSE (Root Mean Square Error).

The following subsections discussed the models used in this study:

GARCH Models

The ARMA models are used to estimate the mean model and ARCH/GARCH-type models are used to model the volatility as a solution to heteroscedasticity. Engle (1982) suggested an ARCH model and Bollerslev (1986) suggested the GARCH model. When the variance of errors in a time series follows moving average in squared errors, the model is called an ARCH model, if the variance of errors follows an autoregressive moving average (ARMA) model the model is a GARCH model (autoregressive conditional heteroscedasticity. The ARMA(p,q) -GARCH *(s,.r)* mean model can be written as:

$$X_t = \psi_t + \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_p X_{t-p} + \cdots + \vartheta_1 \epsilon_{t-1} + \cdots + \vartheta_q \epsilon_{t-q} + \epsilon_t$$

$$\epsilon_t / \Psi_{t-1} \sim N(0, \sigma_t^2)$$

variance model

$$\sigma_t^2 = \varpi + \sum_{i=1}^{s} \alpha_i \epsilon_{t-i}^2 + \sum_{j=1}^{r} \beta_{\mathrm{j}} \sigma_{t-j}^2$$

Where α_i is the parameter of the ARCH parameter and β_{j} is the GARCH parameter and the variance process is positive and stationary.

TGARCH Models:

The threshold GARCH model (TGARCH) developed by Glosten, Jagannthan, and Runkle in (1993) and Zakoian (1994). The ARMA(p, q)-TGARCH (s, r) mean model is specified as:

$$X_t = \psi_0 + \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_p X_{t-p} + \cdots + \vartheta_1 \epsilon_{t-1} + \cdots + \vartheta_q \epsilon_{t-q} + \epsilon_t$$

variance model as

$$\sigma_t^2 = \varpi + \sum_{j=1}^{r} \beta_{\mathrm{j}} \sigma_{t-j}^2 + \sum_{i=1}^{s} \alpha_i \epsilon_{t-i}^2 + \sum_{l=1}^{u} \gamma_l \epsilon_{t-l}^2 \, \Gamma_{t-l}$$

Where Γ_t=1 if ϵ_t<0 and 0 else.

EGARCH Model

Nelson (1991) introduced the EGARCH(Exponential generalized autoregressive heteroscedastic) model that addresses the inefficiencies of GARCH handling of time series data. According for asymmetric effects among positively and negatively distributed for instance. The ARMA(p, q)-EGARCH (p, q) mean model specification is as follows.

$$X_t = \psi_t + \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_p X_{t-p} + \epsilon_t + \cdots + \vartheta_1 \epsilon_{t-1} + \cdots + \vartheta_q \epsilon_{t-q}$$

variance model

$$ln\sigma_t^2 = \varpi + \sum_{i=1}^{s} \alpha_i \left|\frac{\epsilon_i}{\sigma_{t-i}}\right| + \sum_{j=1}^{r} \beta_j \, ln\sigma_{t-j}^2 + \sum_{l=1}^{u} \gamma_l \frac{\epsilon_i}{\sigma_{t-l}}$$

PARCH Model

The PARCH (Power Generalized Autoregressive Conditional Heteroscedastic) model was proposed by Ding, Granger and Engle(1993). In this model, the standard deviation is determined rather than the variance. The PARCH mean model can be stated as follows:

$$X_t = \psi_t + \psi_1 X_{t-1} + \psi_2 X_{t-2} + \psi_p X_{t-p} + \cdots + \vartheta_1 \epsilon_{t-1} + \cdots + \vartheta_q \epsilon_{t-q} + \epsilon_t$$

Variance moder

$$\sigma_t^\eta = \varpi + \sum_{j=1}^{r} \beta_j \, \sigma_{t-j}^\eta + \sum_{i=1}^{s} \alpha_i \left(|\varepsilon_{t-i}| - \gamma_i \varepsilon_{t-i}\right)^\eta$$

Where η the power parameter of standard deviation.

Distribution Assumptions

In this study, three distributions are applied with the GARCH models, Normal, Student-t and Generalized Error Distributions to assess estimation and forecasting performance of these models.

Normal Distribution

The Normal distribution also known as the Gaussian distribution is commonly used to estimate GARCH models. For normal distribution, the log-likelihood function is.

$$lnL(\vartheta_t) = -\frac{1}{2} ln[2\pi] - \frac{1}{2} \sum_{t=1}^{T} ln(\sigma_t^2) - \frac{1}{2} \sum_{t=1}^{T} \frac{\varepsilon_t^2}{\sigma_t^2}$$

where ε_t is the residuals

T is the total number of observations and t is time.

Student-t Distribution

The t distribution is called Student-t distribution and was introduced in the GARCH model framework by Bollerslev (1987)..

$$lnL(\vartheta_t) = -\frac{1}{2} ln\left[\frac{\pi\,[\nu-2]\,\Gamma\left[\frac{\nu}{2}\right]^2}{\Gamma\left[\frac{\nu+1}{2}\right]^2}\right] - \frac{1}{2} ln\sigma_t^2 - \frac{[\nu+1]}{2} ln\left[1 + \frac{[\varepsilon_t]^2}{\sigma_t^2[\nu-2]}\right]$$

Where $\nu > 2$ is the degree of freedom

Γ is the gamma function.

Generalized Error Distribution (GED)

The loglikelihood function of generalized error distribution (GED) is of the form

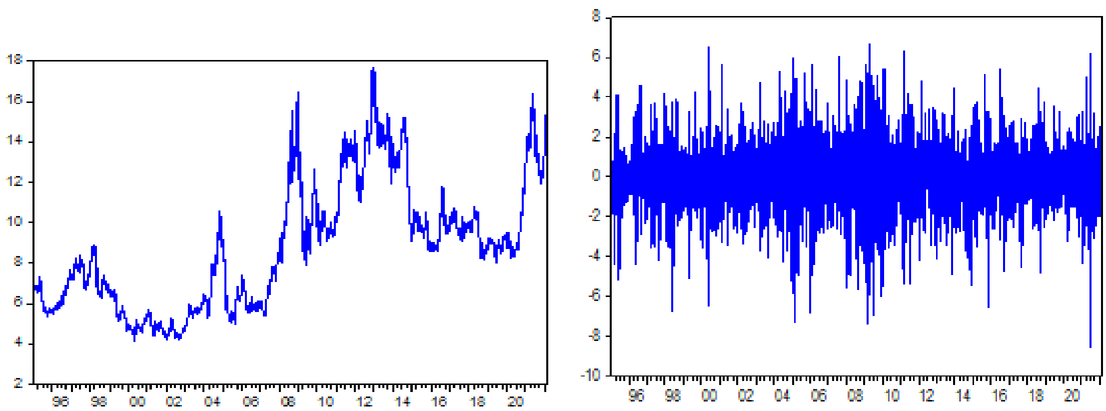
$$L_T(\phi) = \sum_{t=1}^{T}\left[\ln(\frac{\upsilon}{\lambda}) - \frac{1}{2}\left|\frac{\eta_t}{\lambda}\right|^{\upsilon} - (1+\upsilon^{-1})\ln(2) - \ln\Gamma\left(\frac{1}{\upsilon}\right) - \frac{1}{2}\ln(\sigma_t^2)\right].$$

where $\upsilon > 0$ is the degree of freedom and tail-thickness parameter and

$$\lambda = \sqrt{2^{\left(\frac{-2}{\upsilon}\right)}\Gamma\frac{1}{\upsilon}\Gamma\frac{3}{\upsilon}}$$

Results and Discussion

The plots of soybean prices (Y_t) at level are given in Figure 1(A) showing the non-stationary of the series. After taking the 1st log difference Figure 1(B) shows a high variation in data but no trend. These variations refer to the mean-reverting and stationary of the series at the first log difference. The resultant series is denoted as R_t.



(A) Soybean Prices (Y_t) (B) Return Series (R_t)

Figure 1: Plot of Soybean Daily Prices

The descriptive statistics of the soybean prices of Y_t series and the transformed series R_t are given in Table 1. The results for Y_t series show that the mean is greater than median. The coefficient skewness that indicating the distribution is positively skewed. The kurtosis is platykurtic and Jarque-Bera test with p-value <0.005 indicate that the distribution is not normal. The descriptive statistics of R_t series also show the non-normal behavior with high kurtosis.

Table 1
Descriptive Results

Variables	Mean	Median	SD	Skewness	Kurtosis	Jarque-Bera	P-value
Y_t	8.779	8.643	3.1386	0.5260	2.3098	466.623	0.0000

R_t	0.0001	0.0004	0.0142	-0.282	5.6480	2160.224	0.0000

The Augmented Dickey-Fuller (ADF) unit root test is also used to examine the stationarity, and the findings are shown in Table 2. The data process is clearly non-stationary at level and become stationary by taking 1st log difference.

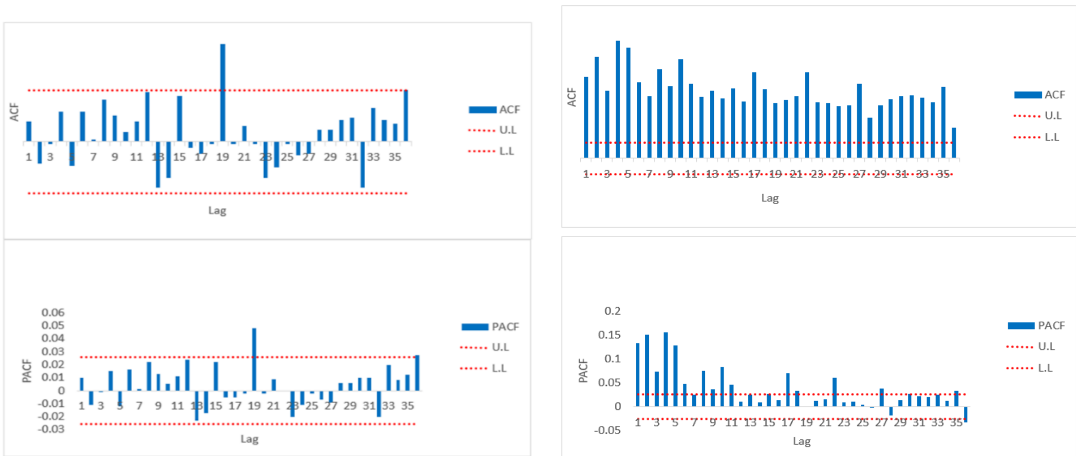
Table 2
Results of Unit Root Test

ADF Test	t-statistic	Probability
Y_t	-2.339879	0.41160
R_t	-82.49995	0.00001

After converting the series in stationary, next we have applied Box and Jenkins (1976) methodology and estimate some tantative ARMA(p, q) models with different orders. We also observed the correlogram of ACF and PACF of the residual and squared residuals of these models. The models satisfying the the diagonostic along with the least values of the AIC, BIC, HQ and log-likelihood are reported in Table 3. Among all of these models, on basis of minimum values of AIC, BIC, HQ and maximum value of Log-Likelihood, ARMA (2, 2) model is select the best model.

Table 3
Model Selection Criteria for ARMA Model

ARMA model	AIC	BIC	HQ	Log Likelihood
ARMA (2,2)	**-5.591883**	**-5.584956**	**-5.589473**	**16132.99**
ARMA (2,3)	-5.5917	-5.58362	-5.58889	16133.47
ARMA (2,4)	-5.591700	-5.583617	-5.588887	16133.46
ARMA (3,2)	-5.591389	-5.582152	-5.588175	16133.57
ARMA (3,3)	5.591207	-5.58015	-5.587591	16134.04
ARMA (3,4)	-5.591645	-5.580098	-5.587627	16136.30



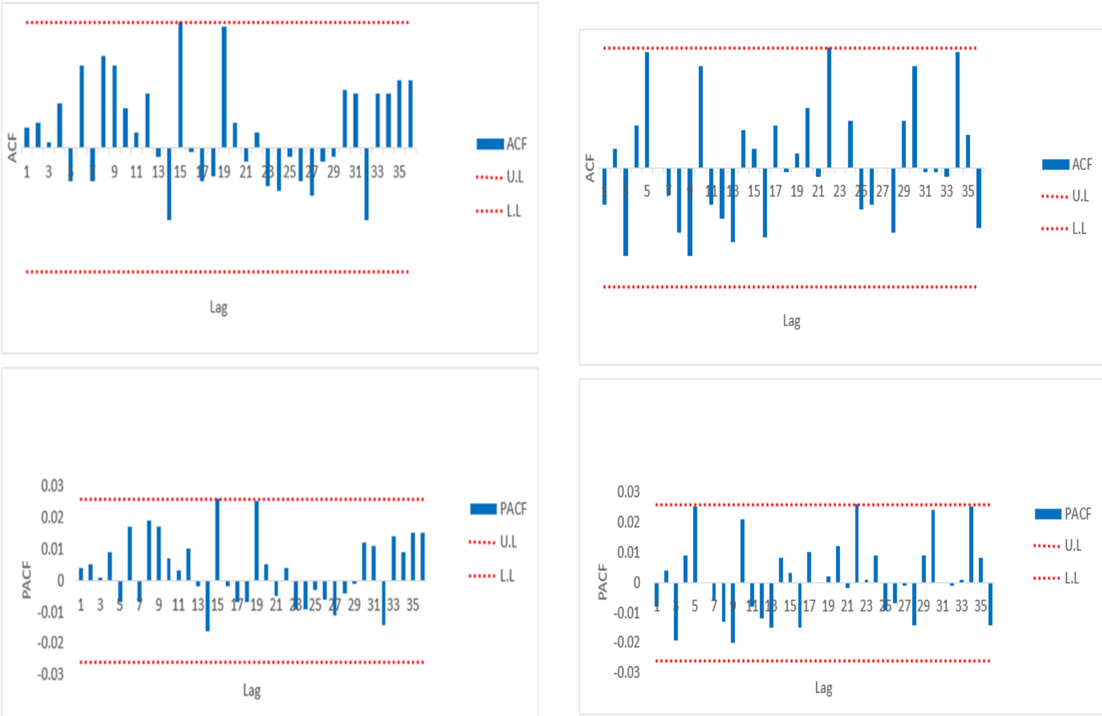
(A) Residuals (B) Squared Residuals

Figure 2: Correlogram of ACF and PACF for ARMA (2,2) Models

As exhibited by Figure 2, the ACF and PACF of the squared residuals are not within limits showing the problem of conditional heteroscedasticity. The GARCH models are used for the solution of this problem. The model identification of the GARCH model with Normal distribution, based on the ACF and PACF values. A number of ARMA models with GARCH models have been fitted to the data and the diagnostics have been checked, some of the models satisfying diagnostics along with the values of AIC, BIC, HQ and Log Likelihood values are shown in Table 4. On the basis of these results ARMA (2, 3)-GARCH (1, 1) model is selected as the best model with the lowest values of AIC, BIC, HQ and highest value of Log Likelihood. The diagnostics of the model show no heteroscedasticity in the model which is clearly seen in Figures 3 (A) - (B).

Table 4
Model Selection for ARMA-GARCH Model with Normal Distribution.

	AIC	BIC	HQ	Log Likelihood
ARMA (2,2)- GARCH (1,1)	-5.7474	-5.7382	-5.74423	16583.6
ARMA (2,3)- GARCH (1,1)	**-5.749220**	**-5.738828**	**-5.745604**	**16589.75**
ARMA (3,2)- GARCH (1,1)	-5.748214	-5.737822	-5.744598	16586.85
ARMA (3,3)- GARCH (1,1)	5.748877	-5.737330	-5.744859	16589.7
ARMA (3,4)- GARCH (1,1)	-5.748804	-5.736103	-5.744384	16590.5
ARMA (4,4)- GARCH (1,1)	-5.749263	-5.735407	-5.744442	16592.8



(A) Residuals (B) Squared Residuals

Figure 3: Correlogram of ACF and PACF for ARMA (2, 3)-GARCH (1, 1)

The similar process of model identification has been repeated for the ARMA-GARCH models with student-t distribution. Some of the models satisfying the diagnostics along with the values of AIC, BIC, HQ and loglikelihood function are reported in Table 5

Table 5
Model Selection for ARMA-GARCH Model with Student-t Distribution

ARMA-GARCH models	AIC	BIC	HQ	Log Likelihood
ARMA(2,2)GARCH(1,1)	**-5.77830**	**-5.767926**	**-5.774702**	16673.67
ARMA(2,3)GARCH(1,1)	-5.778113	-5.766567	-5.774096	16674.08
ARMA(3,2)GARCH(1,1)	-5.778114	-5.766567	-5.774096	16674.08
ARMA(3,3)GARCH(1,1)	-5.777955	-5.765254	-5.773535	16674.62
ARMA(4,4)GARCH(1,1)	-5.777271	-5.762261	-5.772048	**16674.65**

Table 5 represents the ARMA (2, 2)-GARCH (1, 1) is a best model having a lowest value of AIC, BIC, and HQ.

For the model identification of mean ARMA-GARCH model with generalized error distribution, a lot number of models have been estimated and diagnostics are checked but no model satisfied the diagnostics of the residuals. so we exclude the results of these models and considered only two distributions, normal and student-t distribution.

In real life situations, due to some political issues, natural disasters and many other factors the variance process response differently to positive and negative news. To account for negative and positive asymmetries in the models, asymmetric GARCH models such as TGARCH, EGARCH, and PARCH with the selected densities are applied in this study. The estimated results of the selected models satisfying the diagnostics are reported in Table 6. These results reveal that all the ARMA parameters, ARCH parameters, GARCH parameters and asymmetric parameters are highly significant indicating that the conditional heteroscedasticity and asymmetric effect matter and successfully capture by the selected modeling techniques.

Table 6
Estimation Results

	Normal Distribution				**Student t-Distribution**			
Mean Model	**ARMA (2, 3)-**				**ARMA (2, 2)-**			
Coefficients/Variance Model	**GARCH (1, 1)**	**TGARCH (1, 1, 1)**	**EGARCH (1, 1, 1)**	**PARCH (1, 1, 1)**	**GARCH (1, 1)**	**TGARCH (1, 1, 1)**	**EGARCH (1, 2, 1)**	**PARCH (1, 1, 1)**
ψ_o	0.000102 (0.000173)	0.000217 (0.000176)	0.000289 (0.000172)	0.00026 (0.000174)	0.000233 (0.000163)	0.000299 (0.000163)	0.000314* (0.000161)	0.000319* (0.000159)
ψ_1	1.01605** (0.049752)	1.010194** (0.047539)	1.020647** (0.049751)	1.017501** (0.050449)	1.036818** (0.043115)	1.035947** (0.043481)	1.034772** (0.000161)	0.754981** (0.020996)
ψ_2	-0.90202** (0.045922)	-0.90676** (0.043954)	-0.89576** (0.04542)	-0.89958** (0.046183)	-0.91263** (0.042039)	-0.91268** (0.042447)	-0.9093** (0.041604)	-0.95419** (0.020996)
ϑ_1	-1.00143** (0.052361)	-0.99418** (0.000159)	-1.00795** (0.052081)	-1.00256** (0.052947)	-1.03310** (0.039613)	-1.03231** (0.039971)	-1.0327** (0.038712)	-0.76531** (0.017888)
ϑ_2	0.907509** (0.044539)	0.909699** (0.04282)	0.903184** (0.044113)	0.904447** (0.044773)	0.926192** (0.038398)	0.926134** (0.038774)	0.924579** (0.037704)	0.965762** (0.017757)
ϑ_3	0.009429** (0.015303)	0.012094** (0.015383)	0.00965** (0.015101)	0.010755** (0.015369)	-	-	-	-
ϖ	0.00000** (0.00000)	0.0000 (0.0000)	0.23157** (0.021213)	0.0000 (0.0000)	0.00000 (0.00000)	0.00000 (0.0000)	-0.20054** (0.028414)	0.0000 (0.00000)
α_1	0.06369**	0.082292**	0.14806** (0.00899	0.073471** (0.005293)	0.05905** (0.006184)	0.074054** (0.008442)	0.136422** (0.012179)	0.068446** (0.007001)

	(0.00460)	(0.006544)	7)					
β_1	0.92333** (0.00550)	0.926261** (0.005242)	0.026932** (0.005264)	0.928095** (0.004914)	0.931182** (0.00711)	0.934126** (0.00675)	0.063983** (0.021511)	0.935624** (0.006448)
γ_1	-	-0.03654** (0.006907)	0.986232** (0.002077)	-0.17404** (0.033301)	-	-0.03094** (0.009014)	-0.03711* (0.021646)	-0.16417** (0.048278)
γ_2	-	-	-	-	-	-	0.988769** (0.002802)	0.002802
η	-	-	-	1.39621** (0.147663)	-	-	-	1.346809** (0.209795)

Note: Standard errors are given in parentheses; "**" and "*" = the significance at 1% and 5% level of significance respectively;

The forecasts on the basis of all the models given in Table 5 are obtained. The forecasting performance of selected models is compared using the measures MSE, RMSE, and MAPE. The forecast evaluation results of selected symmetric and asymmetric GARCH models with normal distribution are shown in Table 4. The ARMA (2, 3)-TARCH (1, 1, 1) model shows the lower value of RMSE and MAE. Therefore, this model is selected as the best model to forecast the selected series.

Table 6
Forecast Evaluation for Normal Distribution

Models	Root. MSE	MA .Error	MAP. Error
ARMA(2,3)-GARCH (1,1)	0.011622	0.008500	**104.1282**
ARMA(2,3)-TARCH (1,1,1)	**0.011619**	**0.008494**	104.3920
ARMA(2,3)-EGARCH (1,1,1)	0.011629	0.008502	105.5739
ARMA(2,3)-PARCH (1,1,1)	0.011624	0.008497	104.9371

The forecasting performance of the selected models with student-t distribution is given in Table 7. The empirical resuuts favour ARMA(2,2)-TARCH (1,1,1) as the best forecasting modelwith student-t distribution.

Table 5
Forecasting Evaluation Student-t distribution

Models	Root .MSE	MA .Error	MAP. Error
ARMA(2,2)-GARCH (1,1)	0.011624	0.008499	104.3827
ARMA(2,2)-TARCH (1,1,1)	**0.011623**	**0.008497**	**104.6524**
ARMA(2,2)-EGARCH (1,2,1)	0.011628	0.008501	105.0923
ARMA(2,2)-PARCH (1,1,1)	0.011626	0.008499	104.9184

For the selection of final model, the comparison of forecast evaluation of the selected models under Normal and Student-t distribution, the comparison is given in Table 8.

It is obvious that the forecasting performance for this series is best for ARMA(2,2)-TGARCH (1,1,1) model with Student- t distribution.

Table 8
Forecasting Evaluation for Normal and Student-t Distribution

Distribution	Models	Root .MSE	MA .Error	MAP. Error
Normal	ARMA(2,3)-TGARCH (1,1,1)	0.011623	0.008497	104.6524
Student-t	**ARMA(2,2)-TGARCH (1,1,1)**	**0.011619**	**0.008494**	**104.3920**

Conclusion

In this study, the forecasting ability of symmetric and asymmetric GARCH models with normal and non-normal (t-student and GED) innovation for soybean daily prices in Pakistan are compared. The basic analysis of the selected data set reveals non-stationary and conditional heteroscedasticity. ARMA models are applied and find out the problem of heteroscedasticity. To handle this problem the GARCH models are applied along with ARMA models and observed that these models successfully solved the problem of heteroscedasticity. Empirical results also reveal that the GED distribution is not appropriate for this data as no model satisfied the diagnostics of the model in term of ACF and PACF of residuals and squared residuals. So the comparison of the forecasts is based on only the two distributions, normal and student-t-distribution. The forecasting performance of the selected models with normal distribution leads to conclude that TARCH model performed best as compare to GARCH, PAECH, and EGARCH models. Under the assumption of t-distribution, among the symmetric and asymmetric GARCH models, TARCH model also performs the best as compared to other models. The comparison of the model on the basis of distributions, the empirical results favour TARCH model with student t- distribution. Finally, on the basis of the empirical results, it is concluded that the asymmetric GARCH models along with the t-student distribution represent best forecasting ability as compared to symmetric GARCH models and the Student-t distribution is more efficient than normal and GED distributions in modeling and forecasting the daily Soybean prices in Pakistan.

Recommendations

On the basis of this research, we highly recommend that

- Symmetric and asymmetric GARCH models in the other field of life may be promoted due to the importance of these models in capturing features of financial time series, which is not possible with simple ARMA models.
- Accurate forecasts of soybean prices plays a crucial role in policymaking decisions for various stakeholders in different sectors, including farmers, traders, consumers, investors, supply chain managers, and policymakers to navigate the uncertainty and optimize risks for better operational efficiency. As soybean prices are highly influenced by a wide range of factors, including weather conditions, global supply and demand dynamics, government policies, and international trade, reliable forecasting is essential for navigating price volatility and anticipating market trends. Given the complex nature of these influencing factors, advanced forecasting models capturing the structural change such as regime switching GARCH models, states pace models time series modeling approach should be applied to enhance accuracy and reduce uncertainties in price forecasting.
- Soybean is an important oilseed with increasing demand from the feed and chemical industries, including the use of soybean oil and meals as a raw material in the industry, which in Pakistan has increased significantly but in Pakistan, this crop has been neglected because of the area under cultivation declining and the absence of varieties with high yield potential. The government should take steps to develop new high-yielding, latest production technology and strategies that promote domestic economy and farmer training while discouraging heavy imports from North and South America, the agriculture industry will increase seed exports while satisfying fundamental soybean demand.

References

Abraham, E. R., Mendes dos Reis, J. G., Vendrametto, O., Oliveira Costa Neto, P. L. D., Carlo Toloi, R., Souza, A. E. D., & Oliveira Morais, M. D. (2020). Time series prediction with Artificial Neural Networks: an analysis using Brazilian soybean production. *Agriculture*, *10*(10), 475.

Arı, M., & Tüfekci, S. (2014). A comparison of volatility models in forecasting soybean prices. *Agricultural Economics Review*, *15(2), 85-101.*

Bollerslev, T. (1986). Generalized Autoregressive Conditional Heteroscedasticity. *Journal of Econometrics*, *31, 307-327*.

Chi, Y. N. (2021). Time Series Forecasting of Global Price of Soybeans using a Hybrid SARIMA and NARNN Model: Time Series Forecasting of Global Price of Soybeans. *Data Science: Journal of Computing and Applied Informatics*, *5(2), 85-101.*

Ding Z., Granger, C.W.J. and Engle, R.F., 1993. A long memory property of stock market returns and a new model. *Journal of Empirical Finance*, 1, *83-106*.

Felipe, I., Mól, A., & Almeida, V. (2012). Application of ARIMA models in soybean series of prices in the north of Paraná. *Custos e@ gronegócio Online*, *8*.

Gökbulut, R. I., & Pekkaya, M. (2014). Estimating and forecasting volatility of financial markets using asymmetric GARCH models: An application on Turkish financial markets. *International Journal of Economics and Finance*, 6(4), 23-35.

Lence, S. H., & Hayes, D. J. (2002). Forecasting agricultural commodity price volatility using EGARCH models. *American Journal of Agricultural Economics, 84(4), 879-894.*

McMillan, D., & Speight, A. (2004). Forecasting stock market volatility: A comparison of the GARCH and EGARCH models. *Journal of Financial Econometrics, 2(2), 159-185.*

Nelson, D. B. (1991). Conditional heteroskedasticity in asset returns: A new approach. *Econometrica: Journal of the Econometric Society, 347-370*.

Pasha, G. R., Qasim, T., & Aslam, M. (2007). Estimating and forecasting volatility of financial time series in Pakistan with GARCH-type models. *The Labore Journal of Economics , 12(2) 115-149*

Peters, J. P. (2001). Estimating and forecasting volatility of stock indices using asymmetric GARCH models and (Skewed) Student-t densities. *Preprint, University of Liege, Belgium, 3(19-34), 2.*

Shamiri, A., & Hassan, A. (2007). Modeling and Forecasting Volatility of the Malaysian and the Singaporean stock indices using Asymmetric GARCH models and Non-normal Densities. *Journal of Malaysian Mathematical Sciences, 1, 83-102.*

Sharma, S., Aggarwal, V. and Yadav, M.P. (2021). Comparison of linear and non-linear GARCH models for forecasting volatility of select emerging countries. *Journal of Advances in Management Research, 18*(4), 526-547

Tinker, J. N., Gerlow, M. E., Irwin, S. H., & Zulauf, C. R. (1989). *Forecasting Soybean Complex Prices: Univariate and Multivariate Time Series Models.*

Wiles, P. S., & Enke, D. (2015). Optimizing MACD parameters via genetic algorithms for soybean futures. *Procedia Computer Science*, *61, 85-91.*

Xu, W. A. (2011). Determinants of the International Soybean Prices. *Available at SSRN 1734046.*

Xu, Y., Xia, Z., Wang, C., Gong, W., Liu, X., & Su, X. (2021). An Empirical Analysis of the Price Volatility Characteristics of China's Soybean Futures Market Based on ARIMA-GJR-GARCH Model. *Journal of Mathematics*, *2021(1), 7765325.*

Zakoian, J. M. (1994). Threshold heteroskedastic models. *Journal of Economic Dynamics and Control, 18(5), 931-955*