Multivariate Time Series Modelling Of Ex-Pump Prices Of Petroleum Products In Ghana

Chapter 4: Results and Discussions

Group 41

Kwame Nkrumah University of Science and Technology

March 5, 2024

Table of Contents

Introduction

Objective

The purpose of the study is to obtain a suitable model for the ex-pump prices of petroleum products in Ghana.

To examines how changes in the prices of one product cause changes in the price of others in both the short and long run.

Data spanning January, 2007 to June, 2015 are obtained from the National Petroleum Authority of Ghana, covering four petroleum products; Gasoline, Gasoil, Kerosene, and Liquefied Petroleum Gas (LPG) .

Chapter 4: Result And Discussion

This chapter analyses and discusses the results. It presents results of the association between the prices of the products considered, namely;

- Gasoil
- Gasolin
- Kerosene
- Liquefied Petroleum Gas (LPG)

All associated tests and models are generated with R

RoadMap

Start Up

- ◇ Plotting and Descriptive Statistics
- ◇ Stationarity Test
- ◇ Differencing If Not Stationary
- ◇ Plotting of ACF and PACF

RoadMap

Estimation Of Model

- ◇ Lag Length Selection (LLS)
- ◇ Cointegration Test
- ◇ Long Run Equilibrium
- ◇ Short Run Equilibrium
- ◇ Estimation of VEC Model (If There is cointegration)
- ◇ Model Validation
- ◇ Forecast of Ex-Pump prices of Products

Descriptive Statistics

In all, 204 observations are used (January, 2007 to June, 2015).

Training data of 144 observations (January 2007 to December 2012) for modelling

Testing data of 60 data points (January 2013 to June 2015) for model validations.

The descriptive statistics of the products are shown in Table 1 on page 8

Summary Statistics

Table: 1 Summary Statistics

Statistics	GASOIL	GASOLINE	KEROSENE	LPG
Mean	122.445	123.570	82.989	94.766
Maximum	175.480	177.090	120.420	136.190
Minimum	11.600	49.170	6.470	58.500
Standard Deviation	32.306	31.817	27.186	20.609
Skewness	-0.201	0.1307	-1.988	0.413
Kurtosis	3.374	2.123	6.293	2.292
Number of Observations	144	144	144	144

● Highest ● Lowest

Plot of Original Series

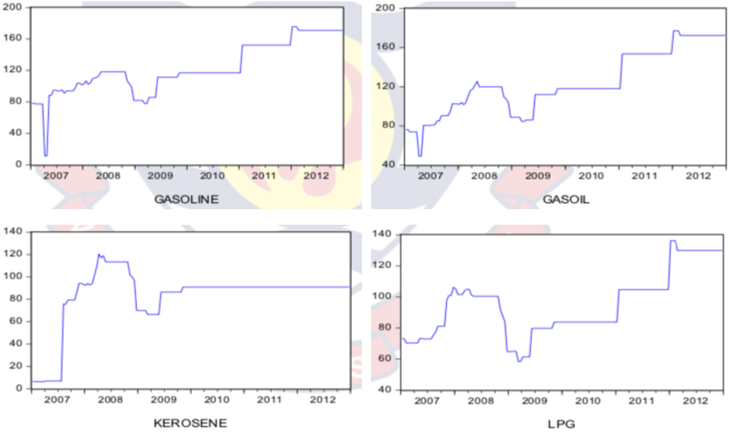


Figure: 1 Time Series Plot of the Original Series

Staionarity Test

We have numerous ways of testing for the presence of a unit root. We have chosen to apply

Augmented Dickey-Fuller Test

$$H0: \text{The series is not stationary}$$
$$H1: \text{The series is stationary.}$$

Phillips-Perron Unit Root Test

$$H0: \text{The series is not stationary}$$
$$H1: \text{The series is statrionary.}$$

KPSS Test for Level Stationarity

$$H0: \text{The series is stationary}$$
$$H1: \text{The series is not statrionary.}$$

Stationarity of Original Series

Table: 2 Univariate URTs of the Original Series

		(Test Statistics)		(P-Values)	
Series	Lag Order	ADF	KPSS	ADF	KPSS
GASOLINE	5	-2.738	2.370	0.269	0.010
GASOIL	5	-2.450	2.437	0.389	0.010
KEROSENE	5	-3.106	0.709	0.166	0.010
LPG	5	-1.975	1.497	0.587	0.010

Is Stationary ?

It is observed that for ADF, all the p-values of the series are greater than 0.05 and this indicates non stationarity. The KPSS test also showed the same results. We now difference the series since the series are not stationary.

Differencing

First Difference

Since all the series (Gasoline, Gasoil, Kerosene, LPG) are not stationary we perform 1st differencing in order to achieve stationarity; The figure 2 below is a plot after the first differencing .

Plot of First Differenced Series

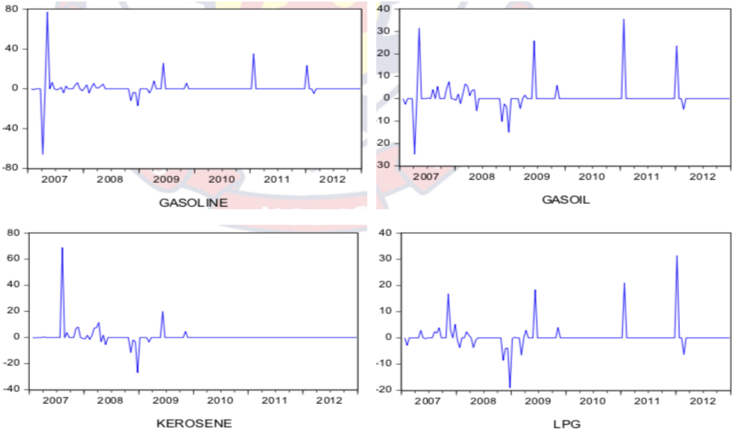


Figure: 2 Time Series Plot of the Original Series

Stationarity of First Differenced Series

Table: 3 Univariate URTs of the Differenced Series

		(Test Statistics)		(P-Values)	
Series	Lag Order	ADF	KPSS	ADF	KPSS
GASOLINE	5	-7.781	0.031	0.010	0.10
GASOIL	5	-5.537	0.045	0.010	0.10
KEROSENE	5	-4.493	0.263	0.010	0.10
LPG	5	-4.473	0.063	0.010	0.10

Is Stationary ?

It is observed that for ADF, all the p-values of the series are less than 0.05 and this indicates the stationarity. The KPSS test also showed the same results. We now estimate the models since the series have attained stationarity.

ACF Plot of First Differenced Series

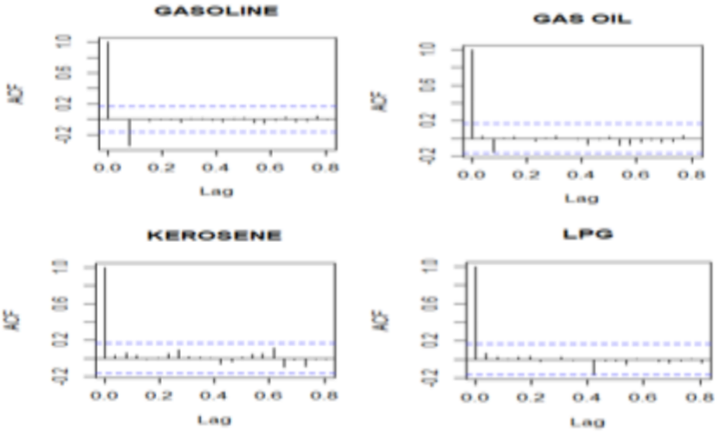


Figure: 3 ACF of the Differenced Series

PACF Plot of First Differenced Series

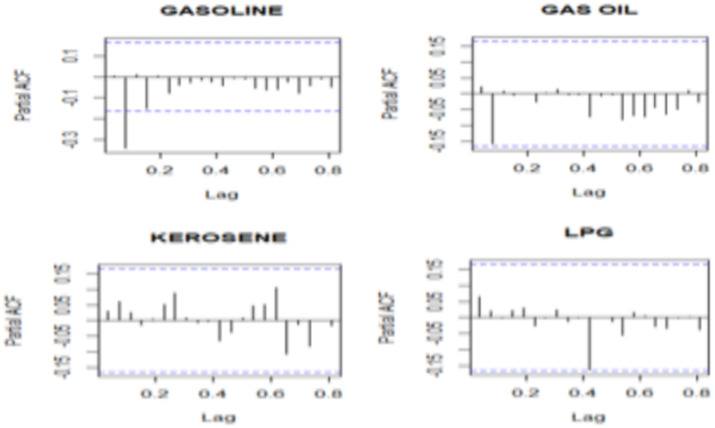


Figure: 4 PACF of the Differenced Series

What Next After Series is Stationary ?

Estimation of VAR/ VEC Models

- ◇ Lag Length Selection (LLS)
- ◇ Cointegration Test
- ◇ Long Run Equilibrium
- ◇ Short Run Equilibrium
- ◇ Estimation of VEC Model (If There is cointegration)
- ◇ Model Validation
- ◇ Forecast of Ex-Pump prices of Products

Estimation of VAR/ VEC Models

Estimating parameters of Vector Autocorelation (VAR) or Vector Error Correlation (VEC) models require that variables are covariance stationary

VAR for instance cannot be used if the variables are not stationary.
Also, if the data is non-stationary, the forecast cannot be done because VAR assumes stationarity

We then test for the long run relationship using Johansen's cointegration test.
That is if the result confirms that there is a long-run relationship among the variables, we can proceed to the VEC model.

The first step involved in estimating is to first determine the lag Length or order.

Lag Length Selection (LLS) Criteria

LLS is significant for VAR/VEC models since selecting too few intervals to result in a cointegrated error and selecting too many intervals may lead to unnecessary loss of degrees of freedom

Three of the LLS criteria are used, namely ;

- FPE (Final Prediction Error)
- AIC (Akaike Information Criterion)
- BIC (Bayesian Information Criterion), aka SC (Schwarz Criterion)

FPE, AIC, and BIC support the inclusion of lag 1 as italicized, and starred in Table 4.

Table: 4 Lag Length Selection Criteria

Lag	FPE	AIC	BIC
0	1.03×10^9	32.107	32.192
1*	117944.1*	23.029*	23.454*
2	127300.7	23.105	23.869
3	142926.7	23.219	24.3224
4	149122.0	23.259	24.701
5	169942.3	23.385	25.167
6	156708.8	23.297	25.419

From Table 4, we can rely on information criteria as only one of these three tests; FPE, AIC, and BIC obtained minimum values at the indicated lag. The test displays lag 1 as the optimum. Thus, the lag length for the estimation is 1.

What Next After Lag Length Selection

Once the unit roots and lag length selections are determined for a time series data, the next step is to inspect whether there exists a Cointegration (Long run relationship) among the variables or not.

Cointegration : Long Run Relationship

Cointegration analysis is important because, if two or more non-stationary variables are cointegrated, a VAR model in the first difference is mis-specified due to the effects of a common trend. The cointegration test determines the type of the regression model to be applied i.e. VAR or VEC

Cointegration Test

H0 : There is no cointegration equation.

H1 : There is a cointegration equation

Table: 5 Determining the Number of Cointegrated Equations

Number of EC	Eigenvalues	Trace Statictics	P-Value	Max-Eigen Staticstics	P-Value
None*	0.358	79.102	0.000	62.959	0.000
At most 1	0.070	16.143	0.702	10.258	0.720
At most 2	0.033	5.885	0.709	4.778	0.769
At most 3	0.008	1.107	0.293	1.107	0.293

Conclusion

Remarkably, the Trace test and max-Eigen statistics suggest the existence of a cointegrated equation (CE).
We shall take into account this fact at the next step.
Since all the series are I(1) and cointegrated, the products ought to be modelled as a VEC model

As a result, a cointegration relationship is obtained.
This throws more light on the long run relationships among the products. Consequently, the products; GASOLINE, GASOIL, KEROSENE, and LPG prices are linked by a long run equation.

Once the unit roots and lag length selections are determined for a time series data, the next step is to inspect whether there exists a long-run equilibrium relationship among the variables or not.

Long Run Relationship

The cointegrating (long-run) relationship is estimated to be;

$$\text{GASOLINE} = -0.0221\ \text{GASOIL} + 0.027\ \text{KEROSENE} - 0.580\ \text{LPG}$$

Thus, with GASOLINE price as the endogenous variable, the long-run relationship indicates that the ex-pump prices of the other products have long run effects.

Specifically, the results indicate that the other products have a negative relation with GASOLINE price in the long run (except KEROSENE), all things being equal.

Log Run Equilibrium

The coefficients of the error correction terms (ECT) [Table 6, 7, 8, 9] show the speed of adjustments of disequilibrium in the period under study.

.

The negative sign associated with the error term is simply a departure in one direction. These are satisfying as they imply convergence in the long run. That is, deviation from the long run is corrected

.

Table: 6 Gasoline Model

Parameters	Coefficient	S.E	t-satistics
GASOLINE Model			
$(\text{Gasoline})_{t\text{-}1}$	0.691	0.189	3.650*
$(\text{Gasoil})_{t\text{-}1}$	-0.0221	0.386	-1.561*
$(\text{Kerosene})_{t\text{-}1}$	0.027	0.091	0.294
$(\text{LPG})_{t\text{-}1}$	-0.580	0.262	-2.211
Constant	0.006	0.805	0.007
ECT	-0.613	0.145	-11.118*

The negative coefficients of the error term for GASOLINE prices indicate dynamic stability, suggesting rapid adjustment speeds.
The magnitude of the correction of the imbalances suggest that , 61.3% of the imbalances in GASOLINE prices are corrected.

Table: 7 Gasoil Model

Parameters	Coefficient	S.E	t-satistics
GASOIL Model			
$(\text{Gasoline})_{t-1}$	0.524	0.126	4.165*
$(\text{Gasoil})_{t-1}$	-0.783	0.256	-3.059*
$(\text{Kerosene})_{t-1}$	-0.002	0.060	-0.030
$(\text{LPG})_{t-1}$	-0.214	0.174	-1.227
Constant	0.017	0.535	0.032
ECT	-0.695	0.096	-7.215*

Concerning GASOIL prices, it indicates 69.5% of shocks in its prices (imbalance) are corrected every two weeks.

Table: 8 Kerosene Model

Parameters	Coefficient	S.E	t-satistics
KEROSENE Model			
$(\text{Gasoline})_{t-1}$	0.058	0.163	0.359
$(\text{Gasoil})_{t-1}$	-0.002	-0.085	-0.256
$(\text{Kerosene})_{t-1}$	-0.518	0.078	-6.652*
$(\text{LPG})_{t-1}$	0.059	0.225	0.263
Constant	0.001	0.692	0.002
ECT	-0.039	0.125	-0.313

For the KEROSENE price, 3.9% of such imbalances are corrected every two weeks

Table: 9 LPG Model

Parameters	Coefficient	S.E	t-satistics
LPG Model			
$(\text{Gasoline})_{t-1}$	0.054	0.106	0.505
$(\text{Gasoil})_{t-1}$	-0.080	0.216	-0.370
$(\text{Kerosene})_{t-1}$	0.010	0.051	-0.197
$(\text{LPG})_{t-1}$	-0.450	0.147	-3.058*
Constant	0.020	0.452	0.044
ECT	-0.036	0.081	-0.437

In the case of LPG price, only 3.6% of such imbalances are corrected.

Short Run Relationship

The short run relationships of the models are explained by the VEC model coefficients as presented in Table 6, 7, 8, 9

Gasoline

Looking at the coefficients, it is observed (Table 6) that in the short-run, GASOLINE price [3.65*] is significant. This is an indication that GASOLINE price exhibits an increment of 69.1% by itself and 2.21% reduction by GASOIL price whiles the others are not significant.

Short Run Relationship

Gasoil

Also, it is observed that GASOIL price [4.17*] is significant by Gasoline. This is an indication that GASOIL price exhibits an increment of 52.4% by GASOLINE price with a 78.3% reduction by itself.

The other products also exhibit both increment and reduction by themselves and/or other products. This is because the coefficients of these products are significant. The short-run results also indicate that the variables influence each other.

Considering GASOLINE price as the dependent variable, it appears the ex-pump prices of the other products influence it. The consequence of this result is that increase ex-pump prices of one or more products are likely to influence others

Long And Short Run Relationship

Now, having analyzed both the short and long-run relationships existing among the variables,
the VEC models are estimated, diagnosed, and validated, and finally, forecasts are generated.

Estimation of VEC Model

The VEC models are estimated using these equations,

$$w_t = c + \phi_{11} w_{t\text{-}1} + \phi_{12} x_{t\text{-}1} + \phi_{13} y_{t\text{-}1} + \phi_{14} z_{t\text{-}1} + \epsilon_t \tag{1}$$

$$x_t = c + \phi_{21} w_{t\text{-}1} + \phi_{22} x_{t\text{-}1} + \phi_{23} y_{t\text{-}1} + \phi_{24} z_{t\text{-}1} + \epsilon_t \tag{2}$$

$$y_t = c + \phi_{31} w_{t\text{-}1} + \phi_{32} x_{t\text{-}1} + \phi_{33} y_{t\text{-}1} + \phi_{34} z_{t\text{-}1} + \epsilon_t \tag{3}$$

$$z_t = c + \phi_{1} w_{t\text{-}1} + \phi_{2} x_{t\text{-}1} + \phi_{3} y_{t\text{-}1} + \phi_{4} z_{t\text{-}1} + \epsilon_t \tag{4}$$

The results of VAR are reported by the 4 equations below. The VEC models are computed with one lag. The models relating the products to their lags and that of others may best be described as;

$$\begin{bmatrix} w_t \\ x_t \\ y_t \\ z_t \end{bmatrix} = \begin{bmatrix} 0.006 \\ 0.017 \\ 0.001 \\ 0.020 \end{bmatrix} +$$

$$\begin{bmatrix} 0.691 & 0.0221 & 0.027 & -0.580 \\ 0.524 & -0.783 & -0.002 & -0.214 \\ 0.058 & -0.002 & -0.518 & 0.059 \\ 0.054 & -0.080 & 0.010 & -0.450 \end{bmatrix} \begin{bmatrix} w_{t\text{-}1} \\ x_{t\text{-}1} \\ y_{t\text{-}1} \\ z_{t\text{-}1} \end{bmatrix} +$$

$$\begin{bmatrix} ECT_{pg} & 0 & 0 & 0 \\ 0 & ECT_g & 0 & 0 \\ 0 & 0 & ECT_k & 0 \\ 0 & 0 & 0 & ECT_l \end{bmatrix} \begin{bmatrix} -1.613 \\ -0.695 \\ -0.039 \\ -0.036 \end{bmatrix}$$

where, $\begin{bmatrix} w_t \\ x_t \\ y_t \\ z_t \end{bmatrix}$, represents the projects GASOLINE, GASOIL, KEROSENE, LPG prices at the time t, $\begin{bmatrix} ECT_{pg} & 0 & 0 & 0 \\ 0 & ECT_g & 0 & 0 \\ 0 & 0 & ECT_k & 0 \\ 0 & 0 & 0 & ECT_l \end{bmatrix}$

refers to the error corrected terms (ECT) for each model, (pg, g, k, and L respectively representing GASOLINE, GASOIL, KEROSENE, and LPG prices), and $\begin{bmatrix} w_{t-1} \\ x_{t-1} \\ y_{t-1} \\ z_{t-1} \end{bmatrix}$ referring to the lags of the products (i.e. lag 1).

The summary of the results of the VEC models is presented in Table 10

Table: 10 Summary Results of the Models

Statistics/Products	GASOLINE	GASOIL	KEROSENE	LPG
F-statistic	31.515	18.810	9.884	8.133
Prob (F-statistic)	0.000	0.000	0.000	0.000
S.E.	9.562	6.351	8.218	5.367
R-squared	53.9%	41.1%	26.8%	23.1%

Table 10 is a summary of the statistics of the VEC models. The results indicate that the models perform creditably well. GASOIL appears to be the best in terms of the variability accounted for.

Model Diagnostics

After modeling, some forecasts are normally estimated. But before the estimated model can be used to generate any forecast,
it is imperative to undertake residual analysis or model diagnostics.

The diagnostic test results include Q- statistics, residual portmanteau test, residual serial correlations, and white heteroscedasticity test.
Tables 11 to 13 provide information on the analysis of the residuals of the models.

Table: 11 VEC Residual Portmanteau Tests for Autocorrelations

Lags	Q-Stat	Prob.	Adj Q-Stat	P-Value	df
1	7.6189	NA*	7.677	NA*	NA*
2	21.603	0.157	21.879	0.147	16
3	29.843	0.576	30.312	0.552	32
4	44.304	0.625	45.228	0.587	48
5	50.472	0.891	51.641	0.867	64
6	111.265	0.012	115.351	0.006	80
7	131.458	0.010	136.684	0.004	96
8	188.126	0.000	197.038	0.000	112
9	197.007	0.000	206.575	0.000	128
10	202.285	0.001	212.289	0.000	144
11	216.121	0.002	227.393	0.000	160
12	230.102	0.004	242.784	0.001	176

VEC Residual Portmanteau Tests for Autocorrelations

The null hypothesis is that there are no residual autocorrelations up to lag h. The test is valid only for lags larger than the selected lag order.
We observe that the residual passes the white noise test since no autocorrelation is left in the VEC model after lag 1.

Lagrange Multiplier Test

The LM (Lagrange Multiplier) test is a statistical test used to access the goodness of fit of a model. It is also know as the Lagrange Multiplier(LM) test for model specification

Table: 12 VECM Residual Serial Correlation LM Tests

Lags	LM-Stat	P-values
1	18.022	0.323
2	18.924	0.273
3	8.073	0.947
4	14.278	0.578
5	6.001	0.988
6	66.795	0.000
7	20.731	0.189
8	69.676	0.000
9	8.623	0.928
10	5.030	0.996
11	14.290	0.577
12	14.019	0.597

Table: 13 VEC Residual Heteroskedasticity

Dependent	R-squared	F (34,163)	P-value	Chi-square (34)	P-Value
res1*res1	0.292	3.415	0.001	38.236	0.290
res2*res2	0.093	0.850	0.615	12.184	0.093
res3*res3	0.216	2.280	0.009	28.266	0.216
res4*res4	0.065	0.573	0.881	8.474	0.065
res2*res1	0.215	2.273	0.009	28.198	0.215
res3*res1	0.155	1.516	0.116	20.262	0.155
res3*res2	0.181	1.835	0.041	23.753	0.181
res4*res1	0.064	0.569	0.884	8.414	0.064
res4*res2	0.049	0.425	0.964	6.387	0.049
res4*res3	0.265	2.980	0.001	34.654	0.264

VEC Residual Heteroskedasticity

Table 13 presents the results of a Heteroscedasticity test, which assesses whether the variances of the error terms in a linear regression model are consistent. The test assumes normal distribution of error terms and aims to verify if variances across the series remain constant.

The findings indicate that the variances are indeed constant, as evidenced by the Chi-Square test's p-value exceeding 0.05. This conclusion is further supported by the p-values from both the F and Chi-square tests.

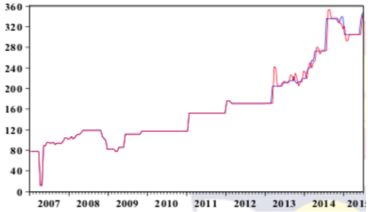
Model Validation

Model validation and certification are crucial in the modeling process, aiding stakeholders in reducing costs, time, and risks associated with extensive product testing. These procedures are essential for establishing trust in statistical models.

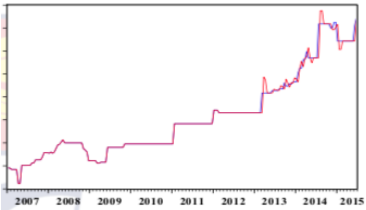
Data was divided into training and validation sets, with 70% of data points used for modeling (2007 to 2012) and the remaining for validation (2013 to 2015). Out-of-sample forecasts for 2013 to 2015 were employed for model validation, comparing predicted product prices to actual prices.

The forecasts closely resembled the original prices, demonstrating the models' effectiveness and accuracy.

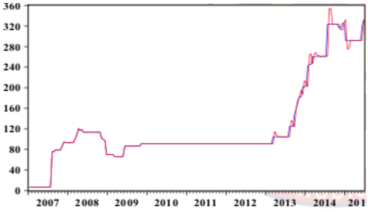
Model Validation



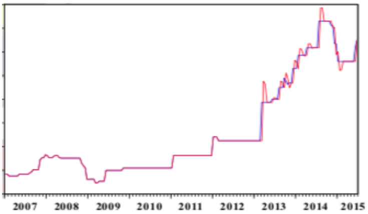
(a) —— GASOLINE (Estimate) —— GASOLINE



(b) —— GASOIL (Estimate) —— GASOIL



(c) —— KEROSENE (Estimate) —— KEROSENE



(d) —— LPG (Estimate) —— LPG

Figure: 5 Model Validation

Forecast of Ex-Pump Prices of Products

After model diagnostics and validation, we, therefore, forecast the prices of the products for the next 12 period as shown in Table 14 on page 47

Table: 14 Forecasts of Ex-pump Prices for the four Products

Period	Premium Gasoline	Gas Oil	Kerosene	LPG
1	172.037	172.360	91.401	130.290
2	173.121	172.360	91.887	130.693
3	173.962	172.360	92.475	131.114
4	174.684	172.360	93.115	131.548
5	175.392	206.830	93.765	131.985
6	176.123	206.830	94.406	132.420
7	176.874	206.830	95.038	132.854
8	177.631	206.830	95.667	133.287
9	178.389	206.830	96.296	133.720
10	179.144	206.830	96.926	134.153
11	179.897	211.110	97.557	134.586
12	180.650	211.110	98.188	135.019

Do forecasted petroleum prices follow the same trend as historical data?

It was noticed that the forecasts for the series consistently rise as time progresses.
This aligns with the trend seen in the original products (Figure 1).
Therefore, the prices of petroleum products have consistently increased over the years, as expected.

Next, we look at causality among the products, hereafter referred to as Granger causality.