Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 6s (2025)

A Robust Regression Analysis of Factors Affecting Diesel and Petrol Price

P.Anandhi[1], Dr. S. Mohan Prabhu[2]

[1]Research Scholar, Department of Statistics, Periyar University Salem – 11, Tamil Nadu, India.

[2]Assistant Professor, and Head, Department of Statistics, Muthayammal College of Arts and Science (Autonomous), Rasipuram, Namakkal, Tamil Nadu-637408, India.

anandhipalanisamy02@gmail.com[1], mohanprabhusu@gmail.com[2]

Article History:

Received: *18-10-2024*

Revised: *02-12-2024*

Accepted: *10-12-2024*

Abstract:

Introduction: A lot of factors influence how much crude oil and gasoline prices fluctuate. Various crises, including war and political unrest, the economic and financial crisis, terrorist actions, and natural catastrophes, have had a significant impact on crude oil and gas prices throughout the last several decades. The robust regression analysis approach is beneficial for investigating the influence of the explanatory variable's diesel and petrol prices. Statistics reveal that some variables, events, and crises have a significant impact on diesel and gasoline prices. To reject the least significant components and proceed to additional regression analysis on the most impacted variables to minimize the number of variables, the ordinary least squares approach was used. The study model explains how these components interact and aids in price forecasting to assist the economy and avoid undesirable scenarios.

Objectives: To identify and understand the key factors influencing the fluctuations in diesel and petrol prices, including supply and demand, crude oil prices, government taxes, and currency exchange rates. To analyse historical and current trends in diesel and petrol prices, enabling predictions about future price movements.

Methods: Robust Regression Analysis for Identify relationships between fuel prices and influencing factors and Correlation Analysis for Measure the degree to which two or more variables. The Trend Charts to use line graphs or bar charts to visualize price changes over time and Forecasting Models to use machine learning algorithms to predict future fuel prices based on historical data.

Results: The following 19 years of gasoline and diesel prices in India served as the foundation for the research project "Forecasting Model for Petrol and Diesel Price" (December 2001–December 2019). To achieve its objectives, this study evaluated secondary data gathered from multiple secondary sources and used a variety of statistical approaches, including trend analysis, regression analysis, and one-way ANOVA. Gasoline shortages are likely to remain, denying fuel to those in most need at any costs, while wealthy countries compete for limited supply.

Conclusions: In summary, the prices of natural gas and crude oil have fluctuated throughout time and will continue to do so. Natural gas and crude oil prices are influenced by a variety of variables, including large environmental disasters, the status of the economy, and the adoption of new technologies. The multiple regression establishes a predictive relationship between variables for both dependent and independent variables. Using data and historical records, the multiple regression technique found and indicated a link between factors that significantly influenced natural gas and crude oil prices.

Keywords: Explanatory variable, Oil and gas, Regression analysis, Residuals, Variance.

1. Introduction

Previous research has shown that supply and demand ultimately determine long-term fluctuations in crude oil and gas prices, whereas short-term fluctuations are influenced by economic and financial factors such as the business cycle and financial market speculation, both of which increase volatility in the oil and gas markets. The relationship between supply and demand influences how volatile the prices of gasoline and crude oil are. The stability of the periphery oil and gas market is dependent on the presence of nations that generate these resources. The price of crude oil and gas has fluctuated significantly over the last few decades because of a variety of unexpected events such as war and political unrest, the economic and financial crisis, terrorist attacks, and natural disasters, all of which amplify the fluctuation of oil and gas prices.

Geopolitics has also had an influence on gasoline and crude oil prices, as the supply is controlled by the major oil producing countries. Conflicts between any of these countries might cause significant problems. If a discrepancy in an oil-producing region looks to be threatening crude stockpiles, gas and oil prices may change. Natural calamities, for example, can have an impact on gasoline and oil prices. These are catastrophes that go beyond the scope of economic theory or management. For example, Hurricane Katrina significantly damaged gas and oil distribution infrastructure along the east coast of the United States in 2005, causing a national energy crisis. Between August 29, 2005, and September 5, 2005, gasoline prices in various areas of the country increased by more than one dollar. Large lineups of people waiting to fill up at petrol stations were caused by a media frenzy over predictions that the fuel supply may soon run out. To keep fuel costs from climbing higher, President George W. Bush withdrew 30 million gallons of gasoline from the US strategic reserves. Hurricane Katrina affected 25% of the US crude oil production. Ten to fifteen percent of the refining capacity was idle during the first several days after the hurricane. A month later, Hurricane Rita caused damage in the Gulf States. The combined effects of the two hurricanes lowered crude oil refinery inputs by 11.7 million barrels per day for the week ending September 30. Since March 1987, this has been the lowest production average.

In addition, the Mississippi River flooding in May 2011 caused at least $2 billion in damage. Fuel prices had climbed to $4.02 per gallon by the second week of the month, owing to supply issues. The damage that floods was creating to oil infrastructure concerned commodity dealers. Unexpectedly, oil spills have little effect on pricing. For example, the Exxon Valdez leaked 11 million gallons (262,000 barrels) of oil. Despite the fact that Alaska's coastline was decimated, the supply and price of oil internationally remained relatively unchanged. The BP oil catastrophe produced 12 times more oil per barrel than the Exxon Valdez. Nonetheless, the price of gasoline and oil has remained relatively unchanged. Why? First, the delayed recovery from the 2008 financial crisis has resulted in a fall in global demand. Second, about 3.4 million barrels of oil, or 134 million gallons, were leaked during a three-month period. Despite the large amount of oil, this accounts for only a small portion of the United States' total oil consumption. The United States claims to have utilized 7.5 billion barrels in 2019. According to the Administration of Energy Records, this is slightly more than 20.5 million barrels a day, or more than six BP oil disasters.

To combat the coronavirus outbreak, some countries began limiting travel and suspending businesses in January 2020. The demand for oil and gasoline began to decline. The average amount of gas and oil consumed in the first quarter of 2020 was 94.4 million b/d, 5.6 million b/d less than at the same time last year [6]. Spending fell until it was negative as storage capacity rose. On April 12, 2020, OPEC and Russia agreed to cut supply to keep prices stable. Since this endeavour failed to persuade traders that supply would not meet demand, gas and oil prices continued to plummet. On April 20, 2020, a barrel of WTI cost around -$37 in Cushing, Oklahoma. Prices, however, immediately recovered, peaking at $40 in the last week of July after falling to $39/b by June 5 and returning to that level in the first week of June. The global average price for Brent crude oil was $43 per barrel in 2020, with $53 per barrel expected in 2021 and 2022. As the need for energy grows, one approach is to switch from fossil fuels to renewables.

1. Methods

2.1 LMS and LTS Estimators

Rousseau developed the Least Trimmed Squares (LTS) and Least Median of Squares (LMS) estimators back in 1984. The LMS approach uses the sample median for summaries because it is a more trustworthy location estimate than the sample mean. This leads to LMS regression, which may be stated as,

$$\min_{b} \operatorname*{med}_{i} (y_i - x'_i b)^2 \tag{1}$$

Rousseau developed the Least Trimmed Squares (LTS) and Least Median of Squares (LMS) estimators back in 1984. The LMS approach uses the sample median for summaries because it is a more trustworthy location estimate than the sample mean. LMS regression, denoted as n-n/3, follows from here. Its asymptotic efficiency against repeated errors approaches zero. There is a closed-form approach for calculating LMS in the location model. In a regression scenario, the LMS could not be estimated originally using a precise technique.

To solve this issue, Rousseeuw and Leroy (1987) improved on the LMS technique by doing a weighted least squares procedure after the first LMS fit. Weights are calculated based on the first LMS fit. Rousseeuw (1987) developed the LTS technique to overcome the inefficiencies of the LMS in this specific context. The LTS approach is used to calculate the parameters from k data points. The coverage value, k, ranges from n/2 to n. The intended objective is:

$$\min_{b} \sum_{i=1}^{k} r^2{}_{[i]} \tag{2}$$

where $r^2{}_{[i]}$, are the ordered squared residuals, k is the trimming constant. The LTS estimator additionally includes an intercept adjustment step. Simply substitute the current intercept with the current LTS regression estimator's "location" LTS estimate of the residuals. The LTS has a high breakdown value of 50% and a high breaking point of (n-(n/2))/n. LTS is more resistant to local variables than LMS. LTS outperforms LMS in terms of statistical effectiveness. The final step in enhancing LTS is to use weighted least squares refinement.

2. Results

We may be interested in estimating (predicting) the value of one variable based on the value of another after examining the link between two variables. The term "dependent" or "explained" refers to the variable predicted because of other conditions, whereas "independent" or "predicting" refers to the other variable. Regression analysis was utilized to generate an average link, which was then statistically assessed as the foundation for the forecast. Depending on whether it is linear or not, the formula is called the explanatory equation or the regression equation. Consider the relationship between two variables, such as rainfall and agricultural production, input cost and product cost, disposable income and consumer spending, and so on.

By illustrating the average relationship between two variables, regression analysis allows for estimation and prediction. When one variable's value is known, the regression statistical approach is used to forecast or estimate the average values of the other variables. The independent variable is explained, but the dependent variable is employed to explain the independent variable.

Case: I

Null hypothesis H_0: There is no correlation between the price of diesel and the price of petrol.

Table 1: Dependent Variable: petrol price Coefficients

Model	SS	Df	MSS	F	Sig.
Regression	4559.761	1	4559.761	48.42	.000[a]
Residual	1601.356	17	94.17		
Total	**6161.117**	**18**			

Predictors: (Constant), demand of crude oil

Table 2: Dependent Variable: petrol price

Model	Un standardized Coefficients		Standardized Coefficients	t	Sig
	B	Std	Beta		
Constant	11.105	5.991	-	1.446	.069
Demand of Crude oil	0.017	0.002	0.868	7.196	.000

Inference

The regression equation for the crude oil is y = a + b(x)

Petrol price = 12.111 + .017 (x) demand of crude oil.

Case -II

Null hypothesis H_0: There is no correlation between the price of diesel and the price of diesel

Table 3: Dependent Variable: demand of crude oil

Model	SS	Df	MSS	F	Sig.
Regression	10141116.010	1	10141116.010	120.211	.000
Residual	1434131.117	17	84360.653		

Total	**11575247.127**	**18**			

Predictors: (Constant), diesel price

Table 4: Dependent Variable: demand of crude oil Coefficients

Model	Unstandardized Coefficients		Standardized Coefficients	t	Sig
	B	**Std**	**Beta**		
Constant	419.567	207.675	-	1.911	.054
Demand of Crude oil	50.621	4.637	.936	10.918	.000

Inference

The regression equation for the demand of crude oil is y = a + b(x)

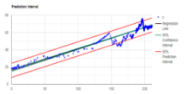
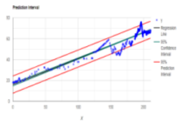
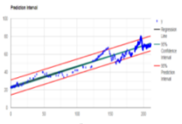
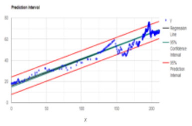
Diesel price = 438.567 + (50.621) x demand of crude oil

Table 5: Dependent Variable: demand of crude oil

Model	SS	Df	MSS	F	Sig.
Regression	6004647	2	3002323	1849.919	.000
Residual	87639	54	1622		
Total	6092286	56			

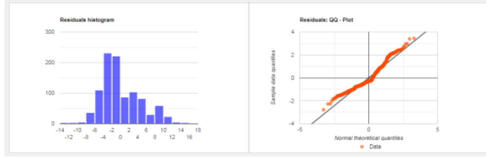
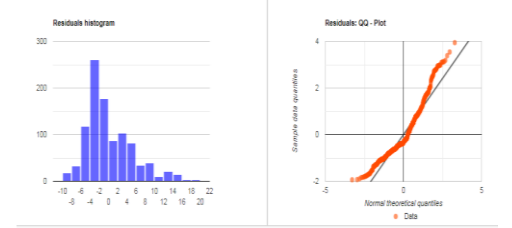
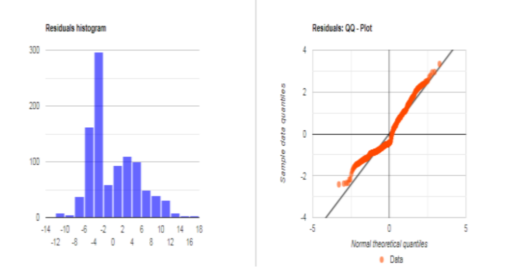
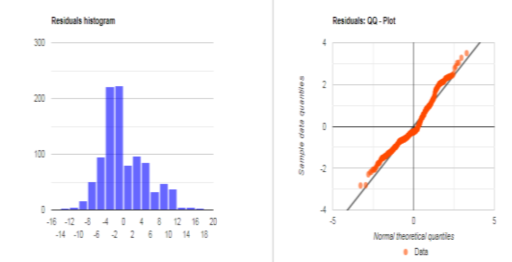
Inference

The calculated p-value (.000) is not statistically significant at the 0.05 level. We consequently conclude that the three oil prices differ considerably and reject our null hypothesis.

States	Estimated Regression Model	Regression chart
Delhi	$\hat{Y} = 15.9846 + 0.3211\ (X)$	
Kolkata	$\hat{Y} = 18.3639 + 0.3802(X)$	
Mumbai	$\hat{Y} = 22.811 + 0.3607(X)$	
Chennai	$\hat{Y} = 17.6928 + 0.4100(X)$	

Source code: pinstat ©

Figure 1: Diesel-Regression fitting model

States	**Estimated Regression Model**	**Regression chart**
Delhi	$\hat{Y} = 33.7826 + 0.2928(X)$	
Kolkata	$\hat{Y} = 38.3281 + 0.269(X)$	
Mumbai	$\hat{Y} = 39.7563 + 0.265(X)$	
Chennai	$\hat{Y} = 37.1162 + 0.276\ (X)$	

Source code: pinstat ©

Figure 2: Petrol-Regression fitting model

3. Discussion

The analysis of factors affecting diesel and petrol prices highlights the complex interplay of various global, regional, and domestic elements. Crude oil prices are the primary determinant of diesel and petrol costs, as they account for a significant portion of the production costs. Global supply-demand dynamics, geopolitical events, and OPEC decisions play a major role in driving crude oil price fluctuations. hangs in fuel prices have widespread implications, affecting household budgets, transportation costs, and overall economic growth. High fuel prices can lead to inflationary pressures across sectors. Finally, the poorest of the poor should not face environmental injustices perpetrated by wealthier nations or have their options limited just because the cleanest energy sources are employed for growth. However, due to resource scarcity and climate change, the poor may face more challenges than others. Growing and caring for the climate and environment should go hand in hand, especially when the cleaner choice is more reliable and affordable in the long run. A comprehensive understanding of the factors affecting diesel and petrol prices equips stakeholders, including governments, businesses, and consumers, to make informed decisions. Strategic planning, robust policies, and investment in alternative energy sources are essential to ensure energy security and economic stability in the face of fluctuating fuel prices.

References

[1]. Agarwal, S. et al., 2002, Birth weight patterns in rural undernourished pregnant women, Indian Pediatrics, Vol. 39, March, p. 17.

[2]. Barone S., Lombardo A. and Tarantino P., 2007, A weighted logistic regression for conjoint analysis and Kansei engineering, Quality and Reliability Engineering International, Vol. 23 Issue 6, pp. 689 – 706.

[3]. Biswas, S.C., et al., 2000, Impact of some biosocial variables on infant and child mortality, Demography India. Vol. 29, No. 2, pp. 211-221.

[4]. Chernick, M., R., 2008, Bootstrap Methods, A Guide for Practitioners and Researchers. Second Edition, Wiley, Inc., New York.

[5]. Ferri C, Flach P., Hernandez-Orallo J, 2002, Learning Decision Trees Using the Area under the ROC Curve, Nineteenth International Conference on Machine Learning (ICML 2002); Morgan Kaufmann; pp. 46-139.

[6]. Hazarika et al., 2000, Differences in blood pressure level and hypertension in three ethnic, Asia-pacific journal of public health, Vol. 12, No. 2, pp. 71-78.

[7]. Ravi.J & et. al., 2017 , A robust measure of pair-wise distance estimation approach: RD-RANSAC, International Journal of Statistics and Applied Mathematics, 2(2): 31-34.

[8]. Ravi.J & et. al., 2018, A New Robust Distance Machine Leaning Approach to Edge and Corner Detection, RD-FAST, Journal of Statistics and Mathematical Engineering, 4(3), 31-44.

[9]. Weber, S., O.; and Michalik G., W., 2008, Incorporating sustainability criteria into credit risk management, Business Strategy and the Environment, Vol. 19 Issue 1, pp. 39 - 50.