FORECASTING PETROLEUM PRODUCT PRICES

Theresa Hallquist, Nancy Hassett, Paula Weir, Energy Information Administration
Theresa Hallquist, EIA, 1000 Independence Avenue SW, Mailstop: EI-431, Washington, DC, 20585

Key Words: Time Series, ARIMA Models, Transfer Function Models

Objective of the Project

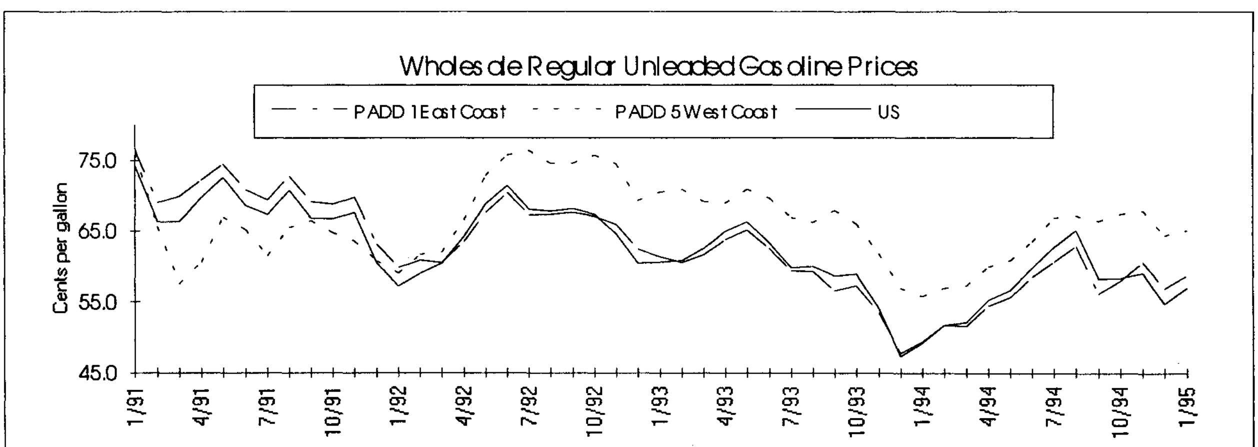
The Energy Information Administration (EIA) wishes to produce price estimates of selected petroleum products in a more timely manner. Currently, estimates for petroleum product prices are published 2-3 months after the end of the reference period. Our customers would like to have the information earlier than we usually publish. Evidence of this was seen in the results of the EIA customer surveys done in 1995 and 1996. The information from the two customer surveys shows high levels of satisfaction with customer service provided by EIA staff. However, the surveys show lower levels of satisfaction with timeliness of our data release relative to the other areas the surveys measured. As a result, EIA has targeted timeliness as an area needing improvement. The goal of this project is to publish prices within 2 weeks of the end of the reference month. The accuracy of the estimates should be within 1 cent. These estimates would then be finalized through the current process and final numbers would be published according to the current schedule.

Approach

ARIMA (Autoregressive integrated moving average) transfer function models were chosen to forecast petroleum product prices. The petroleum product prices are collected on two forms, the EIA-782A, "Refiners' Monthly Petroleum Product Sales Report," and the EIA-782B, "Resellers/Retailers' Monthly Petroleum Product Sales Report," referred to as the EIA-782 price in this article. Transfer function models are ARIMA models which use input data series as predictors. These were appropriate because the petroleum prices were dependent not only on their own past history but also on one or more independent data series. ARIMA models were used because serial correlation between data points is often encountered when using economic time series data. ARIMA models allow for autoregressive or moving average structures.

We used an iterative model-building process:
1. Calculate simple monthly averages for daily input price series.
2. Check each input series for stationarity (constant mean) -- if not stationary, use first differences.
3. Calculate cross-correlations between response series and input series where the cross-correlation is defined as the correlation between the two series at different lags. (Note: Cross-correlation is not meaningful if the input series exhibits auto-correlation. The input series must be pre-whitened. This is done by fitting an ARIMA model to the input series and taking the residuals. Then fit the same model to the response series and take those residuals. Compute the cross-correlations between the two residual data series.)
4. Use the cross-correlation functions to identify the form of the relationship between the response series and the input series.
5. Estimate the transfer function portion of the model and analyze the residuals.
6. Use the residuals to estimate the noise model.

The ARIMA transfer function model consists of two parts: a transfer function which relates the input variables to the EIA-782 price and a noise model which is fit to the residual error series. The final model should not contain any autocorrelation in the residuals nor should there be any correlation between the input variables and the residuals.



The ARIMA transfer function may be expressed as:

$$Y_t = \mu + \Sigma \frac{\omega_i(B)}{\delta_i(B)} B^k X_{it} + \frac{\Theta(B)}{\Phi(B)} a_t$$

$$= \textit{transfer function} + \textit{noise model}$$

where:

Y_t	= EIA-782 price at time t
μ	= mean
$\omega_i(B)$	= transfer function model numerator polynomial (analogous to moving-average operator)
$\delta_i(B)$	= transfer function model denominator polynomial (analogous to autoregressive operator)
B^k	= backshift operator $B^k X_t = X_{t-k}$
X_{it}	= i-th predictor time series at time t
$\Theta(B)$	= moving-average operator
$\Phi(B)$	= autoregressive operator
a_t	= random error

The steps used in generating forecasts were:
1. Estimate model using less data to check for the stability of parameter estimates (typically 12 months less).
2. For each month m, fit model of functional form described above, using data up to and including that month, m. Using those parameter estimates, forecast prices for months m+1, m+2, assuming input series are known for months m+1 and m+2.

Wholesale Regular Unleaded Gasoline

Although many petroleum product forecasting models have been developed, this paper will present results for wholesale regular unleaded gasoline forecasts at the U.S. and Petroleum Administration for Defense District (PADD) level. PADDs are regions of the country, the U.S. is split into 5 PADDs. A national level price estimate was calculated as a volume-weighted average of the PADD level estimates. These volume weights were obtained from another EIA survey which measures prime supplier sales of petroleum products. Original PADD-level models for wholesale regular unleaded gasoline prices used New York Harbor unleaded gasoline spot prices as the sole input variable. Spot prices are prices quoted for immediate delivery of gasoline at a trading center. The models were highly sensitive to large changes in the spot price. The addition of finished motor gasoline ending stocks as another input variable resulted in only minimal improvements in the models. Ending stocks are stocks of gasoline held in storage on the last day of the month.

Industry analysts suggested that rack prices would be a better predictor of EIA-782 prices. Rack prices are truckload or smaller sales of gasoline where title transfers at a terminal. Both rack prices and the West Texas Intermediate (WTI) crude oil spot price were examined as potential additional predictors. Rack prices from January 1991 through July 1995 were obtained for one 'representative' city in each PADD. These cities were: New York City for PADD 1, Chicago for PADD 2, Houston for PADD 3, Denver for PADD 4, Los Angeles for PADD 5.

The input data series that were tested as potential predictors in each PADD were as follows:

PADD	Predictors	Statistically Significant
1	WTI crude oil spot price	X
	New York harbor unleaded gasoline spot price	X
	PADD 1 monthly ending stocks	
	PADD 1 rack price	X
2	WTI crude oil spot price	
	New York harbor unleaded gasoline spot price	
	Gulf Coast unleaded gasoline spot price	X
	PADD 2 monthly ending stocks	
	PADD 2 rack price	X
3	WTI crude oil spot price	
	Gulf Coast unleaded gasoline spot price	X
	PADD 3 monthly ending stocks	
	PADD 3 rack price	X
4	WTI crude oil spot price	X
	Los Angeles unleaded gasoline spot price	X
	Gulf Coast unleaded gasoline spot price	
	PADD 4 monthly ending stocks	
	PADD 4 rack price	X
5	WTI crude oil spot price	X
	Los Angeles unleaded gasoline spot price	
	PADD 5 monthly ending stocks	
	PADD 5 rack price	X

Different formulations of gasoline are required to be sold regionally for environmental reasons. Two examples are reformulated gasoline (RFG) and oxygenated gasoline. To meet the environmental requirements, gasoline may be blended with Methyl Tertiary Butyl Ether (MTBE) or ethanol. Seven different formulations of gasoline appeared in the rack price series: regular unleaded, ethanol blend regular unleaded, ethanol RFG oxygenated regular unleaded, ethanol RFG regular unleaded, MTBE blend regular unleaded, MTBE RFG oxygenated regular unleaded, MTBE RFG regular unleaded. In some cases, more than one formulation appeared for a city. It was unclear which price series should be used. Thus, various scenarios were tested to see which produced the most accurate forecasts. Models were tested and developed under three scenarios:

1. LIKELY -- Using the price of the formulation most likely to be sold during a particular month as determined by industry analysts.
2. RACK -- Using the average of the price series for those formulations sold during a particular month.
3. WTDRACK -- Using a weighted average of the price series for those formulations sold during a particular month, where the prices are weighted by the estimated percentage of the population in that PADD using that type of gasoline.

If prices for both MTBE-based and ethanol-based formulations (e.g. MTBE RFG regular unleaded and ethanol RFG regular unleaded) appeared in the data for a particular month for a city, they were given equal weight in the calculation of the estimated rack prices.

Model building was done iteratively. First, spot prices were added, then ending stocks were added, and finally, rack prices and crude oil spot prices were considered. The rack price was found to be a significant predictor of the wholesale regular unleaded gasoline price in each PADD. The rack price consistently appeared in every model, regardless of how the rack price was derived (LIKELY, RACK, or WTDRACK). The WTI crude oil spot price and the rack price were determined to be significant predictors for the PADD 1 wholesale regular unleaded gasoline price. For PADDs 2 and 3, the Gulf Coast unleaded gasoline spot price and the rack price were both significant. For each of PADDs 1 and 3, the three models were of the same form, with only small differences between the coefficients. For PADD 2, the models using LIKELY and WTDRACK were of the same form with small differences between the coefficients. For PADD 4, the WTI price and the rack price were significant predictors. In addition, the Los Angeles spot price was also significant when using WTDRACK as the estimator of the rack price. For both PADDs 4 and 5, the best models were obtained using the actual rack prices rather than the first differences of the rack prices. For PADD 5, the two models based on the LIKELY rack price and the WTDRACK price contained terms for both the WTI crude oil price and the rack price. The coefficients for these two models were similar in magnitude. In contrast, the model based on RACK did not depend on the WTI crude oil price.

Actual minus Forecast (cents)						
	Added Spot		Added Ending Stocks		Added Rack	
Month	1 month ahead	2 month ahead	1 month ahead	2 month ahead	1 month ahead	2 month ahead
1	3.9	7.0	3.4	6.3	-.1	.8
2	2.2	2.4	2.2	2.1	.9	-.3
3	-.1	2.5	.1	2.9	-1.2	0
4	2.5	3.1	2.6	3.2	1.2	1.6
5	-.5	1.7	-.5	1.4	.4	.4
6	1.9	3.4	1.2	2.6	-.1	.3
7	1.0	2.8	.9	2.8	.4	3.7
8	1.5	-5.3	1.5	-5.2	3.4	3.1
9	-7.1	-5.3	-6.6	-4.6	-.3	-.3
10	3.7	-7.1	3.6	6.9	0	1.3
11	2.4	-1.3	3.0	-1.0	1.3	1.8
12	-3.9	-1.3	-3.5	-.4	.3	.4
Avg. Abs. Diff.	2.6	3.6	2.4	3.3	.8	1.2

After input data series were added, forecasts were generated for one year, beginning with forecasts made in June 1994. One-month ahead and two-month ahead forecasts were compared against actual values as a means of measuring the performance of the models. As expected, the one-month ahead forecasts were more accurate than the two-month ahead forecasts.

Similar results were obtained under each of the three scenarios for estimating rack prices. For PADDs 1 and 2, 11 of 12 months exhibited differences of 2 cents or less for the one-month ahead forecasts. For the two-month ahead forecasts, 10 of 12 months had differences of 2 cents or less. PADD 1 had a larger number of forecast errors of 1 cent or less than did PADD 2. For PADD 3, both the LIKELY and RACK models resulted in 12 of 12 months having differences of 2 cents or less for the one-month ahead forecasts. For the two-month ahead forecasts, all three models resulted in 10 of 12 months having forecast errors of 2 cents or less. Both the LIKELY and WTDRACK models for PADD 4 gave one-month ahead forecast errors of 2 cents or less for 11 of 12 months. For the two-month ahead forecasts, the RACK and WTDRACK models gave forecast errors of 2 cents or less for 9 of 12 months.

MODEL EQUATIONS

EIA-782 Only: $Y'_t = (1+.86B)/(1-.51B)e_t$

Added Spot: $Y'_t = (.64+0.40B-.09B^3)(1-.29B+.35B^2)S'_t + 1/(1+0.21B^2)e_t$

Added Ending Stocks: $Y'_t = (0.62+0.05B+.09B^2)/(1-^4 0.31B+.15B^4)S'_t + (.06+.07B)/(1+.88B)D + (1-.25B-.43B^2)e_t$

Added Rack: $Y'_t = (0.09+0.63B)/(1+0.36B)C'_t + 0.48/(1-0.36B)S'_t + (.18-0.19B)/(1-0.41B)R'_t + e_t$

where:

Y_t	= EIA-782 price
C_t	= average WTI crude oil spot price
S_t	= average monthly New York Harbor spot price for wholesale unleaded gasoline
D_t	= monthly ending stocks
R_t	= average monthly rack price (RACK)
B	= backshift operator such that $B^k(X_t) = X_{t-k}$
e_t	= random error

and a superscript ′ denotes a first-difference ($X'_t = X_t - X_{t-1}$).

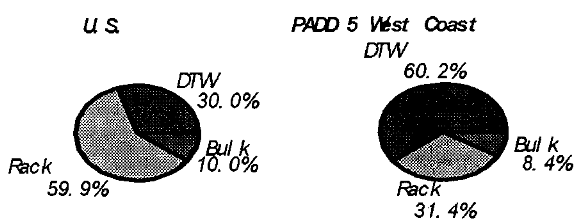
None of the three scenarios resulted in significantly better forecasts than the others. One reason for this may be that for many months, there was no difference between the 3 scenarios. (For example, in Chicago and Houston, there was only one formulation reported from January 1991 through November 1994; so for this time period for these cities, RACK=LIKELY=WTDRACK). Overall, the WTDRACK models did result in slightly more one-month ahead and two-month ahead forecast errors within 1 cent or less. However, the LIKELY models resulted in slightly more one-month ahead forecast errors within 2 cents or less. The RACK and WTDRACK models resulted in slightly more two-month ahead forecast errors within 2 cents or less than did the LIKELY model.

Current models are limited by the types of gasoline sold in that PADD's "representative" city. For instance, in New York City, the representative city for PADD 1, no conventional gasoline has been sold since November 1994, and yet conventional gasoline makes up 59% sales in that PADD. The RFG rack price for PADD 1 may not be a true representation. Also, the breakdown between ethanol-based and MTBE-based gasoline formulations is not known. For modeling purposes, each was given equal weight in the calculation of the rack prices.

One-month ahead and two-month ahead forecasts were generated for one year, starting with June 1994. The differences between the actual EIA-782 prices and the forecast prices were compared. The models performed well except for the time period when RFG was first introduced. The one-month ahead forecasts fell within 2 cents of the actual values with the exception of forecasts made during December 1994 and January and March 1995. The two-month ahead forecasts fell within 2.2 cents with the exception of those forecasts made from December 1994 through March 1995. Those forecasts differed by as much as 5 cents.

PADD 5 West Coast

The model did not fit as well in PADD 5. PADD 5 is the only PADD in which Dealer Tank Wagon (DTW), rather than rack makes up most of the wholesale market for gasoline. DTW sales are sales of gasoline priced on a delivered basis to a retail outlet. Since the DTW price is the largest component of wholesale regular gasoline for PADD 5, it was added to the forecast model for that PADD. The question became: Will the DTW price make a difference if used as an input variable for the PADD 5 forecast? (Buying the DTW data will cost a lot of money.)



To answer this question, 4 models were developed using DTW prices. However, DTW prices could only be obtained from January 1991 through August 1994. This is approximately one year less data than used in the other models.

Model 1 includes the DTW price and the EIA-782 price;
Model 2 includes the EIA-782 price but not the DTW price;
Model 3 includes the DTW price and the first difference of the EIA-782 price;
Model 4 includes the first difference of the EIA-782 price but not the DTW price.

The monthly rack price was estimated as the simple average of the prices for all formulations sold in that month (regular, ethanol blend regular, and MTBE blend regular). This was earlier referred to as the RACK model.

When the DTW price was available, it served as a statistically significant predictor. The DTW price and the lag 1 DTW price were significant predictors in model 1. The rack price and the lag 1 rack price were significant in both models 1 and 2. For those models in which the DTW price did not appear (models 2 and 4), the WTI crude oil price was also a significant predictor. An interesting finding was that for model 4, the WTI crude oil and spot prices were the only significant predictors. Past results suggest that the rack price should have been a significant predictor. For longer data series, that was true. However, for this particular time period of 44 months from January 1991 through August 1994, the spot price appears to be a better predictor than the rack price.

Actual minus Forecast (cents)				
	Model 4 Without DTW		Model 3 With DTW	
Month	1 month ahead	2 month ahead	1 month ahead	2 month ahead
1	-.1	0	-.4	.2
2	.1	.2	.6	1.1
3	0	.6	.5	1.4
4	.6	-.5	.8	.8
5	-1.1	-.9	-.1	.5
6	.2	.3	.6	.8
7	0	-1.3	.1	-.6
8	-1.3	-1.3	-.8	-.4
9	.1	-1.6	.3	-.2
10	-1.7	-1.2	-.5	-.7
11	.6	1.5	-.2	-.1
12	.8	1.1	.2	-.5
Average Absolute Difference	.6	.9	.4	.6

The models in which the first difference of the EIA-782 prices served as the dependent variable provided the best forecasts. Model 3, which includes the DTW price, clearly outperforms the other models. The maximum absolute difference between the actual and forecast prices was .8 cents for the one-month ahead forecasts and 1.4 cents for the two-month ahead forecasts.

The previous models, PADD 5 models included, did not perform well during the time period when RFG was first introduced in January 1995. DTW prices were not available during the time period when RFG was first introduced. Thus, although current results indicate that DTW is a good predictor for the time period from January 1991 through August 1994, one cannot say for sure whether the addition of DTW to the models would necessarily produce more accurate forecasts for subsequent time periods.

Future Enhancements for the U.S. and PADD Models

One improvement for the models would be the addition of rack price data for other cities in order to obtain a more representative PADD-level rack price. Additionally, information on the breakdown between ethanol-based and MTBE-based formulations could improve the monthly rack price estimates.

The overall goal of this project is to publish prices within 2 weeks of the end of the reference month within 1 cent accuracy. The models for wholesale gasoline do achieve this goal. This will result in increased customer satisfaction with respect to timeliness of published estimates.