

Estimating Weekly Stocks of Other Oils

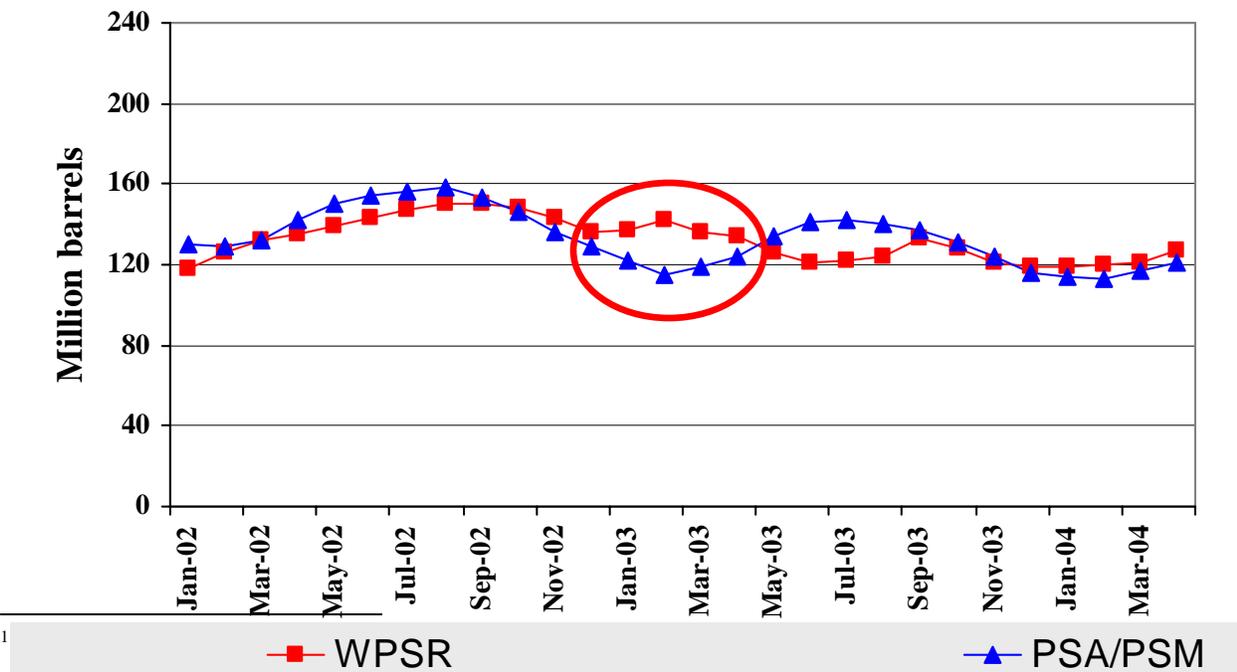
This is a working document prepared by the Energy Information Administration (EIA) in order to solicit advice and comment on statistical matters from the American Statistical Association Committee on Energy Statistics. This topic will be discussed at EIA's fall 2006, meeting with the Committee to be held October 5 and 6, 2006.

Challenge: EIA collects data on *stocks of other oils*¹ on a monthly basis. However, weekly *stocks of other oils* are estimated from the monthly *total petroleum stocks* and *petroleum products supplied* because *stocks of other oils* are not collected on a weekly basis. The results of this data (among others) appear in the Weekly Petroleum Status Report (WPSR). In February 2003, a notable difference was observed between EIA's weekly estimates and the monthly data for the *stocks of other oils* product. (See graph 1). Thus, we began to explore other ways to estimate weekly *stocks of other oils*. Note: The estimation procedure excluded the weekly propane data. We evaluated alternative methods by comparing the weekly estimates with the *Petroleum Supply Monthly (PSM)* data. The *PSM* data used for estimation purposes include weekly data for major oil products, monthly data for major oil products and the "*Stocks of Other Oils*" products.

Comparison of WPSR Estimates of Other Oils Stocks (excluding propane) and PSM Data

Graph 1

January 2002- April 2004



¹

While in some periods the results of the comparisons between the estimated weekly *stocks of other oils* correspond closely with the monthly data, the current procedures are not documented and, therefore, the estimation methodology cannot be replicated. Based on what SMG did, could the committee provide input on other alternative methods to improve the forecast of weekly other oils stocks?

What SMG Did:

SMG investigated other statistical methods for estimating the WPSR *stocks of other oils* with the goal being 1) to obtain a replicable method and 2) to obtain a method that produced estimates that corresponded closely with the PSM data. Methods included *multiple regression* and *unobserved components* models.

The following table and graph show the models/methods attempted and a summary measure of the difference between the estimates obtained from the models and the data collected from the PSM from January 2002 to December 2005. The models were fitted using data from January 1993 to December 2001. In addition, for comparison, differences between the *WPSR* estimates (obtained from the model known as the Weekly Petroleum Supply Report system) and the *PSM* data are also shown.

Table 1.

Model	Variables	Mean Absolute Percent Error 1/2002 to 12/2005
Multiple Regression (Fitted data from January 1993 to December 2001)	Major Petroleum Products and PROPANE and 11 dummy month variables (JAN-NOV) are regressors	4.38% with range [-11.4%, 9.41%]
Multiple Regression (Fitted data from January 1993 to December 2001)	Major Petroleum Products and PROPANE, and 1st difference, 12th difference, 1st & 12th difference of (major petroleum products and PROPANE), and 3 Quarter dummy variables (Q1, Q2, Q3) are regressors	4.05% with range [-8.02%, 12.93%]
Unobserved components	Trend, Season, Cycle, Autoregressive term, A	4.20% with range

model (UCM) (Fitted data from January 1993 to December 2001)	regressive terms involving lagged dependent variables, and a regressive term on independent variables.	[-11.5%, 10.3%]
Multiple Regression (STEO)	Pentanes Plus; Propane; LPGs excluding Propane; Other Hydrocarbons and Oxygenates; and Other Petroleum Products	5.89% with range [-11.9%, 17.8%]
Weekly Petroleum Supply Report System (WPSRS)	19 sets of parameters to choose in a model	5.92% with range [-24.16%, 14.72%]

Graph 2

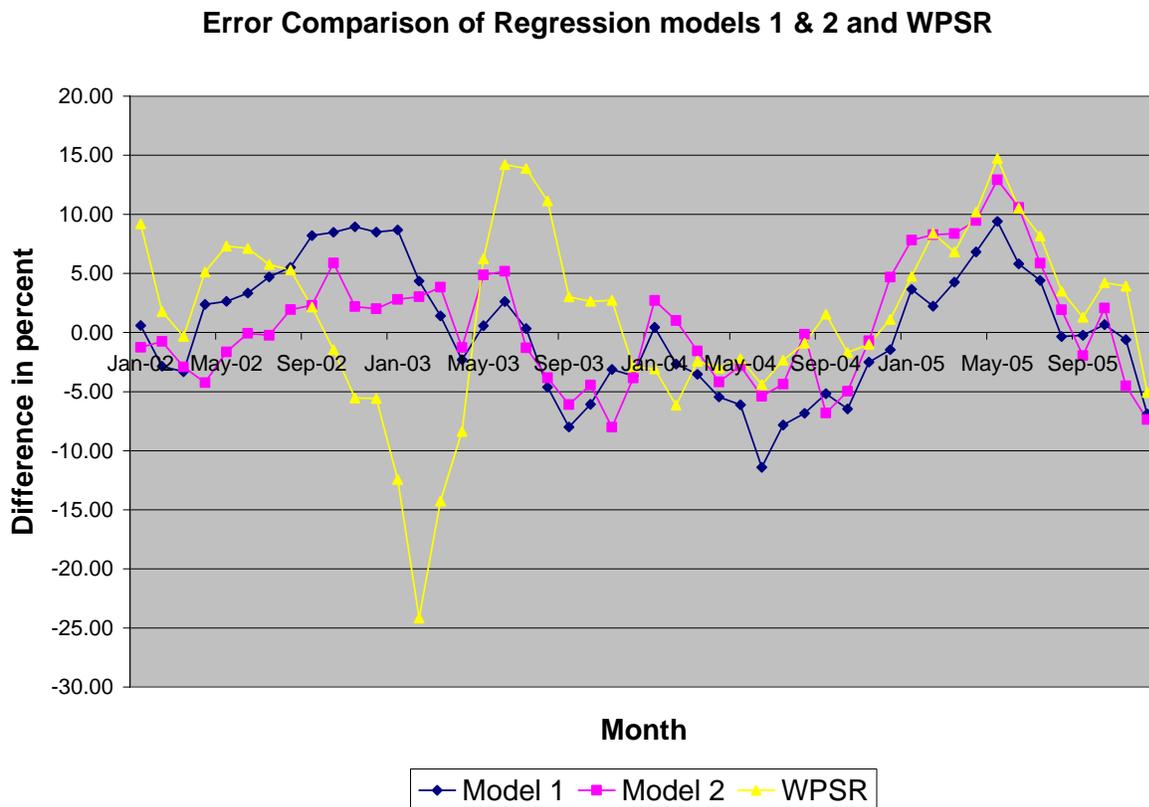


Table 2

Errors	Model 1	Model 2	WPSR
Mean Absolute Percent Error 1/2002 to 12/2002	4.96% with range [-3.32%, 8.94%]	2.12% with range [-4.24%, 5.89%]	4.72% with range [-5.57%, 9.20%]

Mean Absolute Percent Error 1/2003 to 12/2003	3.81% with range [-8.00%, 8.68%]	4.04% with range [-8.02%, 5.19%]	9.67% with range [-24.16%, 14.21%]
Mean Absolute Percent Error 1/2004 to 12/2004	4.99% with range [-11.4%, 0.45%]	3.28% with range [-6.82%, 4.69%]	2.49% with range [-6.16%, 1.56%]
Mean Absolute Percent Error 1/2005 to 12/2005	3.78% with range [-6.90%, 9.41%]	6.76% with range [-7.38%, 12.74%]	6.80% with range [-5.07%, 14.72%]

The above table and graph demonstrated the error rates between the weekly estimated *stocks of other oils* and the *PSM* data from models from January 2003 to April 2004. Although the mean absolute percent error for the *multiple regression* models is within an acceptable range, the regressor variables are based on the monthly major petroleum products stocks data that have a 3 months processing lag and cannot mimic the volatility of weekly data series to reflect the sudden changes of stocks.

The following table and graph show the models attempted, the variables used, and the summary measure (Mean Absolute Percent Error) using the forecasts from the *Unobserved Components Model*, the *Short-Term Energy Outlook (STEO)*, and the *WPSRS*. These three models were used to compare with the data collected from the *PSM* for January 2002 to May 2005. The models were fitted using data from January 1993 to December 2002.

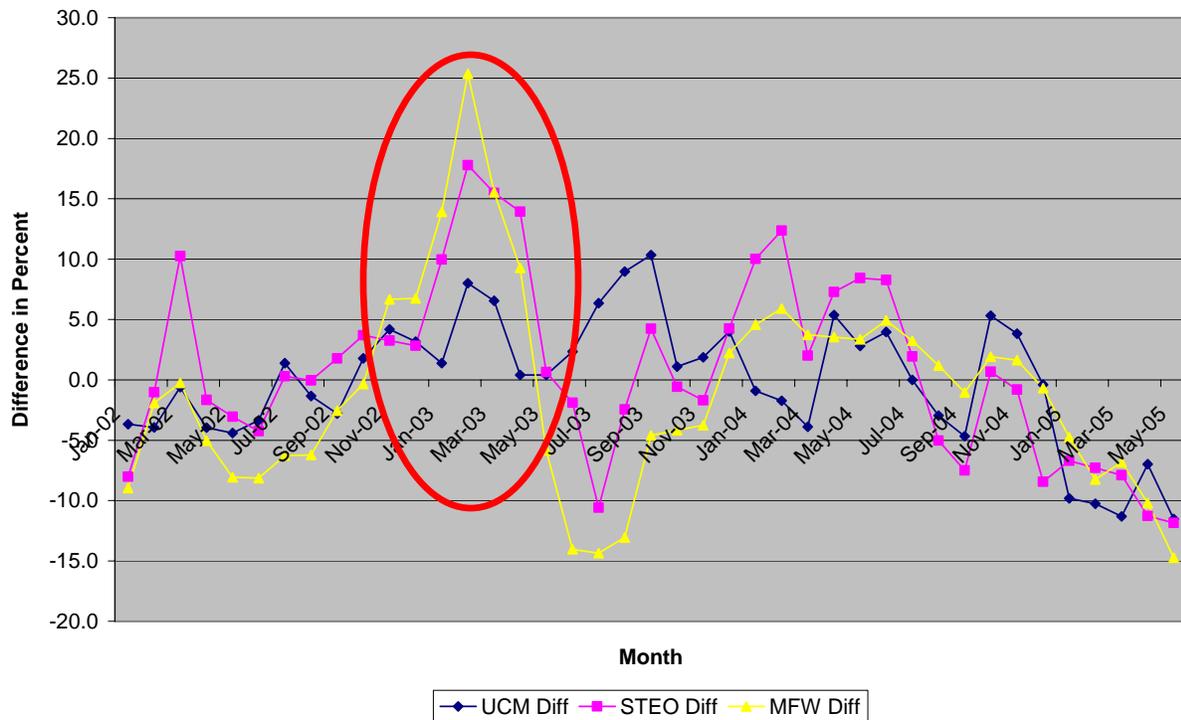
Table 3.

Model	Variables	Mean Absolute Percent Error
Unobserved components model (UCM) (from January 1993 to December 2002)	Trend, Season, Cycle, Autoregressive term, A regressive terms involving lagged dependent variables, and a regressive term on independent variables.	4.20% with range [-11.5%, 10.3%]
Multiple Regression (STEO)	Pentanes Plus; Propane; LPGs excluding Propane; Other Hydrocarbons and Oxygenates; and Other Petroleum Products	5.89% with range [-11.9%, 17.8%]

Weekly Petroleum Supply Report System (WPSRS)	Average daily rate of stock change for the minor products for each month based on monthly data for the past 6 years	6.53% with range [-14.7%, 25.4%]
---	---	----------------------------------

Graph 3

Errors of UCM, STEO and MFW with PSM in Other Oils Stocks



In the third table and graph the *Unobserved Components Model (UCM)* is the most recently developed time series method. The forecasting period covers from January 2002 to May 2005; the mean absolute percent error is within an acceptable range with fewer extreme cases. Moreover, the February 2003 estimate using *UCM* differed by about 7 percent from the PSM estimate compared to 25 percent for the procedure that was used to obtain the WPSR (MFW) data.

STEO estimates *stocks of other oils* data using linear regression equations are estimated using historical monthly data for the following components of *stocks of other oils*: Pentanes

Plus; Propane; LPGs excluding Propane; Other Hydrocarbons and Oxygenates; and Other Petroleum Products.

The dependent variable in each regression equation is the month-to-month change (first difference) in stock levels. The regressor variables may include the deviation in the prior-month stock level from the 4-year average, weather, prices, and monthly dummy variables. The estimated stock changes are added to the latest end-of-month stock published in the Petroleum Supply Monthly to arrive at the stock level forecast.

The forecasting error of UCM is consistently smaller than the error of STEO model and the error of WPSR (MFW) current practice compared with PSM data. UCM can be deployed using the Statistical Analysis System (SAS) easily and the forecasting stocks can be easily documented.

RECOMMENDATION:

1. To collect weekly *stocks of other oils* from the sampled companies. Nine of the largest companies may be contacted to find out how reasonable it would be for them to report one aggregate “other oils stocks” number on a weekly basis. If more than ten need to be contacted, we can conduct a generic clearance to test the feasibility of this collection. Historically, *stocks of other oils* were difficult to collect, but respondents may now have the data in some automated systems.
2. It is not clear as to what is the current practice is to arrive at the weekly *stocks of other oils* number. We would strongly recommend including the unobserved components model (UCM) to be examined by the petroleum products team. We should also reconcile with EMEU to get consensus on how to publish the WPSR and STEO estimates of *stocks of other oils*.
3. Document the procedures/model used to perform the estimation and the rationale for selecting the particular estimate, and make sure that the *stocks of other oils* numbers are reproducible.