

An evaluation of errors in US energy forecasts: 1982–2003

James J. Winebrake*, Denys Sakva

STS/Public Policy Department, Rochester Institute of Technology, 92 Lomb Memorial Dr., Rochester, NY 14623

Available online 30 August 2005

Abstract

Planners, policy-makers, and the private sector rely on energy forecasts to help make policy and investment decisions. In the US, the federal Department of Energy (through the Energy Information Administration and its predecessors) has conducted national forecasts of energy production and consumption for decades. This paper explores US energy forecasts in order to uncover potential systemic errors in US forecasting models. We apply an error decomposition technique to forecasts within each major energy sector (commercial, industrial, residential, and transportation) made during the period 1982 to 2003. We find that low errors for *total* energy consumption are concealing much larger sectoral errors that cancel each other out when aggregated. For example, 5-year forecasts made between 1982 and 1998 demonstrate a *mean percentage error* for total energy consumption of 0.1%. Yet, this hides the fact that the industrial sector was overestimated by an average of 5.9%, and the transportation sector was underestimated by an average of 4.5%. We also find no evidence that forecasts within each sector have improved over the two decades studied here.

© 2005 Elsevier Ltd. All rights reserved.

Keywords: Energy forecast; Forecast error; Energy modeling

1. Introduction

In the US, many energy policy decisions are driven by information and insights derived from energy forecasts. If measured by use in the profession, the “gold standard” for such forecasts seems to be those from the US Department of Energy’s Energy Information Administration (EIA). The EIA has been making energy forecasts for over two decades through its *Annual Energy Outlook* reports.

The importance of these forecasts in defining energy debates cannot be understated. Indeed, at a recent conference on transportation fuels attended by one of the authors, presentation after presentation used EIA forecasts to the year 2025 to justify policy action today. The centrality of these forecasts to energy policy discussions compels us to evaluate such forecasts carefully and understand them in the context with which they are designed. This context is typically one of uncertainty and data gaps that may lead to forecasts

that are highly inaccurate in hindsight. In view of the fact that policy makers rely on forecast results for decision making, the limitations of those forecasts must be transparent and communicated to those policy makers (Munson, 2004). This paper aims to contribute to that transparency.

Previous studies of the accuracy of energy forecasts for the US have indicated limited success in our ability to predict what the energy landscape would look like 20, 10, or even 5 years hence. But these studies have tended to analyze aggregate energy production and consumption. With this paper, we disaggregate and evaluate forecast errors within each major US energy sector (commercial, industrial, residential, and transportation). We identify which sectors experience the greatest forecast errors and whether these errors imply consistent over- or under-estimation of forecasts. We also quantify the contribution of each sector’s forecast error to total forecast error. Finally, we explore whether these forecasts are improving over time. We conduct our analysis using EIA energy forecasts with time horizons ranging from one to ten years made during the period 1982 through 2003. We hope that our results will help

*Corresponding author. Tel: +585 475 4648; fax: 585 475 2510
E-mail address: jjwgpt@rit.edu (J.J. Winebrake).

forecasters improve their forecasts by encouraging a deeper exploration into systemic modeling problems that might be attributed to a particular energy sector.

2. Background

2.1. Energy forecasts in the US

Since 1982 EIA has published energy forecasts in a report called *Annual Energy Outlook* (AEO), the most recent of which at the time of this writing was published in January 2004 (Energy Information Administration, 2004). Between 1982 and 1993 EIA used the Intermediate Future Forecasting System (IFFS) to make projections. The time horizons for these projections varied over the years, with an eight-year forecasting horizon in 1982 to a 17-year horizon in 1993 (Energy Information Administration, 1982–2003). The IFFS was a general equilibrium model that balanced energy supply and demand based on various model inputs, econometric equations, and technology–economic relationships. The IFFS model would produce such output as fuel production and imports, electricity generation, energy prices, and energy consumption by region and by end-use.

In 1994 the IFFS model was replaced by the National Energy Modeling System (NEMS), which is currently used today. NEMS uses the same balanced supply–demand approach as IFFS, but at a more detailed level. NEMS also produces a richer array of output (e.g., greenhouse gas emissions estimates) and allows for a broader evaluation of technology and policy alternatives developed for scenario analysis (Energy Information Administration, 2003b). We rely upon the forecasts from AEO between 1982 and 2003 for our work.

2.2. Previous evaluations of energy forecasts

There are a number of studies in the literature that analyze US energy forecasts. These studies have generally found that from a *numerical* standpoint, forecasting has seen limited success (Smil, 2000, 2003). But different conclusions are drawn to why this is the case. For example, is the forecast wrong because of inaccurate assumptions or poor handling of uncertainty? Or is the forecast wrong because the forecast itself triggered policy action that led to shifts in energy consumption or production?

Ascher (1978) was one of the pioneers of comprehensive analysis of forecasting errors in the energy sector. He studied the accuracy of numerous population, economic, energy, transportation and technological forecasts. Regarding energy demand forecasts, Ascher found no improvement in forecasting over time (Ascher, 1978).

More recently, Vaclav Smil has criticized our dependence upon faulty long-range energy forecasts (Smil, 2000, 2003). Through a comparison of various energy forecasts over a 40 year period, Smil concludes that “long-range forecasters of energy affairs have missed every important shift in the past two generations” (p.176) and “with rare exceptions, medium- and long-range forecasts become largely worthless in a matter of years” (p.124). His analysis shows that almost all long-range US and world energy consumption forecasts from the 1960s through the 1980s were greatly overestimated—by 10% to more than 200%!

The accuracy of EIA projections has been tested by others more systematically. For example, EIA conducts error analyses that can be found in its *Annual Energy Outlook Evaluations* (Holte, 2002; Sanchez, 2002). The last evaluation available at the time of this writing was for 2002 (Sanchez, 2002). In its evaluations, EIA calculates the “average absolute forecast error,” which is expressed as the absolute value of the percentage difference between EIA’s “Reference Case” projections and actual values. This calculation is performed for each year in the forecast. In the most recent report, EIA found errors ranging from -6.7% to 3.8% in their total energy consumption projections.

O’Neill and Desai (2005) applied an error decomposition technique to further analyze AEO errors. Their methodology consists of calculating percentage error, absolute percentage error, mean percentage error, and mean average percentage error for energy consumption, GDP, and energy intensity for all AEOs available. However, their analysis does not disaggregate errors by energy sector. As we will discuss later, the purpose of our work is to explore similar error metrics but for specific energy sectors. The results of their analysis show that low aggregate energy consumption errors were attributed to cancellation effects from larger, oppositely signed errors in projections of GDP and energy intensity (two major drivers behind the NEMS forecasts). Moreover, the authors found no evidence of improvement in forecasting accuracy between 1982 and 2003 (O’Neill and Desai, 2005).

Moving away from US analyses, Linderoth (2002) evaluates forecast errors in IEA countries by calculating forecast errors, average forecast errors, and root mean square of forecast errors for total primary energy consumption, oil consumption, delivered energy (total energy minus losses) by sector, and by IEA country. Linderoth concludes that observed forecast errors were primarily caused by inaccurate growth rate expectations. In addition, he finds that energy consumption in the transportation sector is generally underestimated (Linderoth, 2002). Similar to O’Neill and Desai (2005), Linderoth observes that significant cancellation effects lead to smaller than expected errors in total energy consumption.

Smil (2003) has concluded that no forecasting model, no matter how complex, is able to accurately predict system behavior when extensive social, economic, technical, and environmental interactions exist (as in the energy field). However, others have argued that even with large errors, the forecasting activity is still a worthwhile one. For example, Craig et al. (2002) identify other important purposes of forecasting, such as: a “bookkeeping device”, a communication tool for illustrating complex systems, and an aid to “what-if” analyses. The authors also identify the drawbacks of certain forecasting techniques, pointing out that the use of econometric models (which EIA has used within NEMS) only give reasonable projections when there are no structural changes in the system being modeled; this implies that econometric models are most useful for short-term forecasting since structural changes often do occur in the long-term (Craig et al., 2002).

Despite the differences of opinion on the value of forecasting in the literature, we feel it is important to conduct error analyses that clearly demonstrate where forecast models may be weak. Until now, forecast error studies for US energy consumption only exist on the aggregate level (O’Neill and Desaib, 2005; Sanchez, 2002). Although useful, we think these studies are limited, particularly when applied to an evaluation of NEMS projections, since NEMS includes various sub-models at the disaggregate level. We believe it is important to understand how these forecasts have performed at the disaggregate (energy sector) level. In the next two sections, we explain the methodology we used to examine disaggregated AEO forecast errors and the results of our analysis.

3. Methodology

3.1. Overview

Most energy forecast error analyses in the past have focused on errors in projected total energy production or consumption. One of the questions that remains in such analyses is: *How do each of the major energy sectors contribute to this overall error?* We will demonstrate in this paper that seemingly small errors in total energy forecasts actually hide more significant (but offsetting) errors in specific energy sectors.

In this paper we apply an error decomposition technique to study errors in US energy forecasts by sector (commercial, industrial, transportation, and residential). Because total energy consumption projections from NEMS are additive across energy sectors, forecast errors within each sector will contribute to the overall total forecast error. By breaking down total forecast errors into its disaggregate parts, we can determine what sectors within the NEMS model are

more/less accurate and whether a systemic underestimation/overestimation exists within sectoral model components. We are interested in understanding these errors as related to various forecast time horizons ranging from 1 to 10 years. We are also interested in exploring how forecast accuracy has changed over time; that is, have forecasts become more accurate over the past two decades?

We apply a methodology similar to that found in O’Neill and Desaib (2005). We focus on the “visible error” that they define as “the difference between the projected energy consumption and actual energy consumption.”

In this paper we use two metrics to determine forecast error: *mean percentage error* and *mean absolute percentage error*. Mean percentage error (MPE) is an average error of all forecasts of a given forecast horizon and is given by the function,

$$\text{MPE}_{\tau,j} = \frac{\sum_t (\hat{Y}_{t,\tau,j} - Y_{t,\tau,j})}{n_{\tau,j}} \quad (1)$$

where τ is our forecast horizon (1 year, 2 years, ..., 10 years); t is the year in which AEO was published (1982, 1983, ..., 2003); j is our energy sector index (commercial, industrial, residential, transportation); $\hat{Y}_{t,\tau,j}$ is our forecasted value for period τ and sector j in AEO published in year t ; $Y_{t,\tau,j}$ is our actual value of energy consumption for period τ , sector j , and AEO t ; and $n_{\tau,j}$ is the number of forecasts with time horizon τ for sector j .

MPE calculations for a single forecast horizon (τ) and a single year (t) could take on a positive or negative value. If $\text{MPE} > 0$, then the forecast value was higher than the actual value, and the forecast represents an overestimate. If $\text{MPE} < 0$, then the forecast value was less than the actual value, and the forecast is an underestimate. The reader should note that an average MPE near zero does *not* imply a near perfect forecast. The average may be close to zero, but may represent a combination of highly overestimated and underestimated forecasts that cancel each other out on average.

To more clearly explore the accuracy of forecasts, without concern over whether forecasts are underestimated or overestimated, we apply the mean absolute percentage error (MAPE), given by the following function:

$$\text{MAPE}_{\tau,j} = \frac{\sum_t \frac{|\hat{Y}_{t,\tau,j} - Y_{t,\tau,j}|}{Y_{t,\tau,j}}}{n_{\tau,j}}, \quad (2)$$

where the variables and indices remain the same as in (1). Here, however, we take the absolute value of the error for each forecast, so that the metric is not subject to misinterpretation from cancellation of under- and over-estimated forecasts. We apply both the MPE and

the MAPE on a sector-by-sector basis and in total below.

Both MPE and MAPE identify sector-by-sector forecast errors, but they do not allow for easy consideration of the contribution of these sectoral errors to total error. Because “delivered energy” in the US can be derived as an additive function of energy consumption in all sectors, there is a clear connection between the forecast errors for each sector and the forecast error for energy consumption overall. To determine the contribution of sectoral forecast errors to total forecast error, we introduce the following. Let the total forecast percent error (TFPE_{t,τ}) for a given time horizon τ made at time period t be:

$$TFPE_{t,\tau} = \frac{\sum_j \hat{Y}_{t,\tau,j} - \sum_j Y_{t,\tau,j}}{\sum_j \hat{Y}_{t,\tau,j}} = \frac{\sum_j (\hat{Y}_{t,\tau,j} - Y_{t,\tau,j})}{\sum_j Y_{t,\tau,j}}. \tag{3}$$

Given Eq. (3) we can determine that the contribution of TFPE from a given sector (which we call the *sectoral forecast percentage error* for sector j, or SFPE_{t,τ,j}) is given by

$$SFPE_{t,\tau,j} = \frac{\hat{Y}_{t,\tau,j} - Y_{t,\tau,j}}{\sum_j Y_{t,\tau,j}}. \tag{4}$$

The SFPE_j can also be derived into a *mean sectoral percentage forecast error* (MSFPE_{j,τ}) for a given time horizon τ by modifying (4) as follows:

$$MSFPE_{\tau,j} = \frac{\sum_t \left(\frac{\hat{Y}_{t,\tau,j} - Y_{t,\tau,j}}{\sum_j Y_{t,\tau,j}} \right)}{n_\tau} \tag{5}$$

3.2. Data

For our analysis, we used forecast energy data from the *Supplemental Tables to the Annual Energy Outlook* for 1982–2003. We focus our analysis on the reference case forecasts in these AEOs with the following caveats:

- (1) Before 1990 AEO did not include information about “dispersed renewable energy consumption” (Energy Information Administration, 1998). We follow the approach discussed elsewhere to determine dispersed renewable energy consumption for forecasts made before 1990 (O’Neill and Desai, 2005). Even so, the contributions of dispersed renewable energy consumption in the US tend to be negligible.
- (2) Before 1996 AEO did not report electricity related losses by sector. The “total energy consumption” data by sector in these AEOs is equivalent to what is called “delivered energy” in later AEO versions. Thus, in our analysis we use “total energy consumption” data by sector for pre-1996 AEOs and

“delivered energy” for 1996 and later AEOs. We focus our analysis on delivered energy, since NEMS uses delivered energy to drive other primary fuel use calculations and forecasts.

Actual values for energy consumption by sector were taken from the *Annual Energy Review* (Energy Information Administration, 2003a). It should be noted that actual data for 2003 were taken from the 2004 edition of AEO (Energy Information Administration, 2004) and are considered preliminary data.

We should also note that in this paper we discuss errors of past energy forecasts to measure the accuracy of such projections at a disaggregate level. We do not address the reasons for higher or lower accuracy of certain projections. These variations in accuracy may be a result of different core assumptions and economic conditions, technological and demographical shifts, unaccounted energy substitution between sectors, changes in forecasting methodology, and many other factors.

4. Analysis

4.1. Analysis of MAPE by sector

This analysis offers a closer look at the general accuracy of forecasts, by sector, for time horizons ranging from one to ten years. We can use this analysis to determine if: (a) forecasts exhibit increased uncertainty when time horizons are lengthened; and (b) certain sectors demonstrate a more accurate level of forecasting than others.

Table 1 presents both MAPE and MPE calculations by sector and in total. The results from the MAPE analysis are shown in Fig. 1. Fig. 1 demonstrates that while total energy consumption forecasts have relatively small errors for the time horizons analyzed (ranging from 1.7% to 4.8%), the sectoral errors are much higher. In particular, the transportation sector errors range from a low of 3.0% as an average for 1-year forecasts, to over 11% for 8-, 9-, and 10-year forecasts. The most accurate forecasts seem to be those associated with the residential sector.

4.2. Analysis of MPE by sector

The MPE analysis expands on the MAPE analysis by identifying the directionality of forecast error. We can use this analysis to determine if certain sectors tend to under-estimate or over-estimate forecast errors consistently. This analysis might point to a systemic problem with the forecast models used for a given time horizon. The fact that the MAPE for total energy consumption is lower than errors for individual sectors means that there

Table 1
Energy Consumption Errors by Sector

	Forecast horizon (years)									
	1	2	3	4	5	6	7	8	9	10
Number of observations	14	14	12	11	11	8	6	4	2	3
MPE										
Delivered energy	1.66%	1.58%	1.28%	0.77%	0.10%	-0.94%	-2.53%	-5.40%	-6.74%	-4.86%
Residential	0.79%	0.11%	-0.30%	-1.47%	-0.90%	-3.27%	-2.99%	-0.72%	0.12%	-1.71%
Commercial	-0.44%	-0.50%	-0.69%	-0.88%	-2.31%	-2.47%	-2.07%	-2.40%	-1.87%	-6.61%
Industrial	3.81%	4.84%	5.55%	6.00%	5.88%	5.51%	2.92%	-2.30%	-6.40%	0.41%
Transportation	0.57%	-0.35%	-1.66%	-2.93%	-4.50%	-5.89%	-7.87%	-11.60%	-11.56%	-11.09%
MAPE										
Delivered energy	3.30%	3.64%	3.53%	3.34%	3.20%	3.81%	4.21%	5.40%	6.74%	4.86%
Residential	2.37%	2.54%	2.06%	3.19%	2.83%	4.26%	3.74%	0.98%	1.15%	3.55%
Commercial	2.89%	3.47%	5.13%	5.09%	5.28%	5.35%	4.63%	3.80%	1.87%	6.61%
Industrial	6.32%	7.31%	7.71%	7.70%	8.00%	8.67%	6.86%	3.91%	6.40%	6.08%
Transportation	3.14%	3.80%	5.08%	6.16%	6.63%	7.05%	7.87%	11.60%	11.56%	11.09%
MSPFE										
Delivered energy	1.66%	1.58%	1.28%	0.77%	0.10%	-0.94%	-2.53%	-5.40%	-6.74%	-4.86%
Residential	0.11%	0.00%	-0.06%	-0.25%	-0.15%	-0.53%	-0.49%	10.11%	0.02%	-0.26%
Commercial	-0.05%	-0.06%	-0.09%	-0.12%	-0.27%	-0.30%	-0.25%	-0.27%	-0.20%	-0.73%
Industrial	1.35%	1.72%	1.97%	2.13%	2.09%	1.95%	0.98%	-0.90%	-2.46%	0.11%
Transportation	0.24%	-0.09%	-0.55%	-1.00%	-1.56%	-2.06%	-2.77%	-4.11%	-4.10%	-3.98%

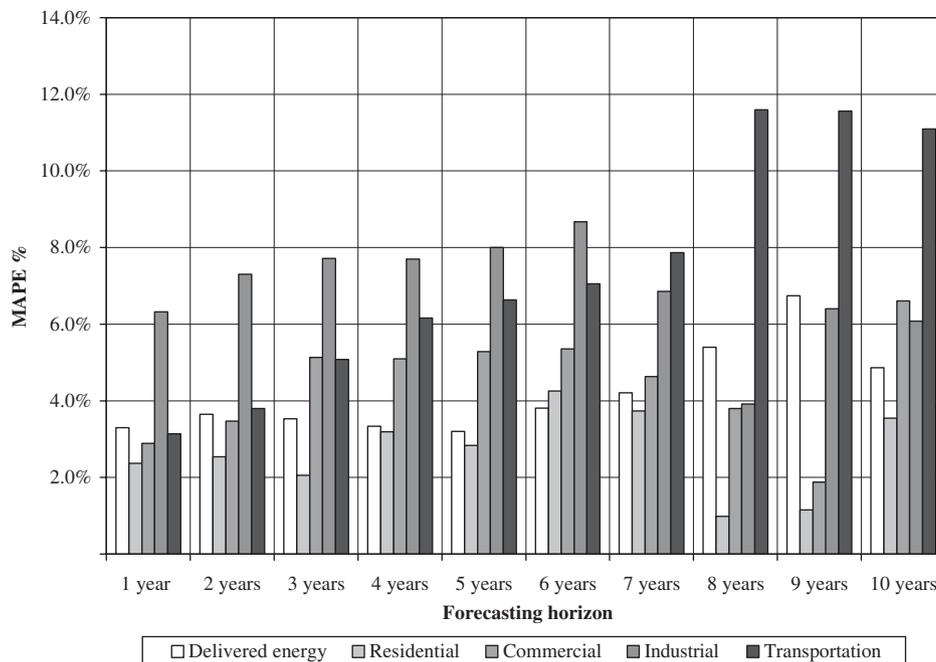


Fig. 1. MAPE for energy consumption by forecast length.

is a cancellation of errors across energy sectors. As previously mentioned, the use of MPE can shed light on the directional aspects of each sector's energy forecasts. MPE calculations are shown in Table 1 and Fig. 2.

For forecasts of five years or less, total energy consumption errors are small on average (around 1%) and are positive (representing overestimation). How-

ever, for forecasts between 6 and 10 years in length, the errors are larger (about 4%) and are negative (representing underestimation). The transportation sector is observed to be highly and systematically underestimated, while the industrial sector tends to be overestimated, particularly for shorter forecast horizons.

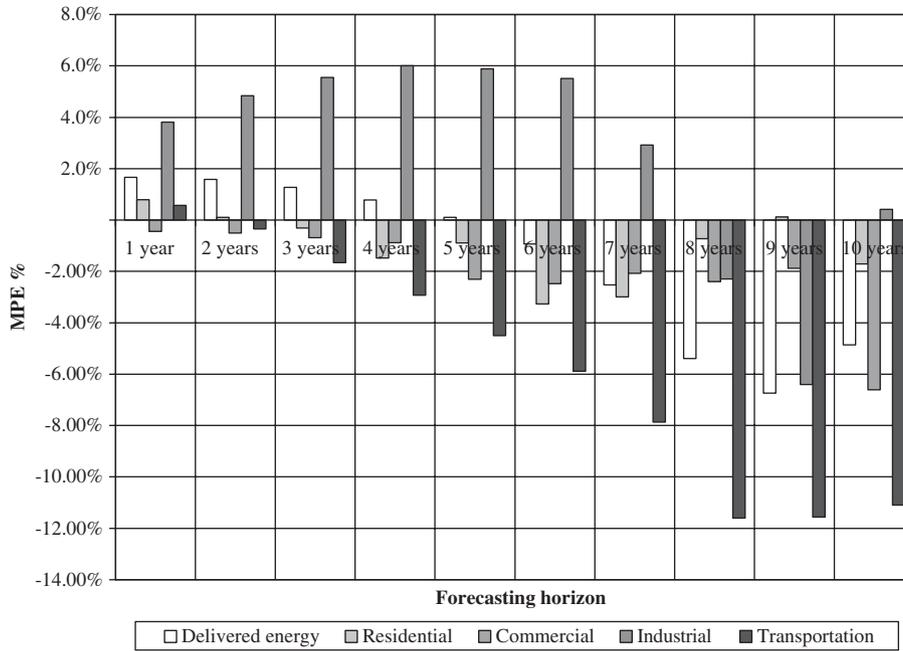


Fig. 2. MPE for energy consumption by forecast length.

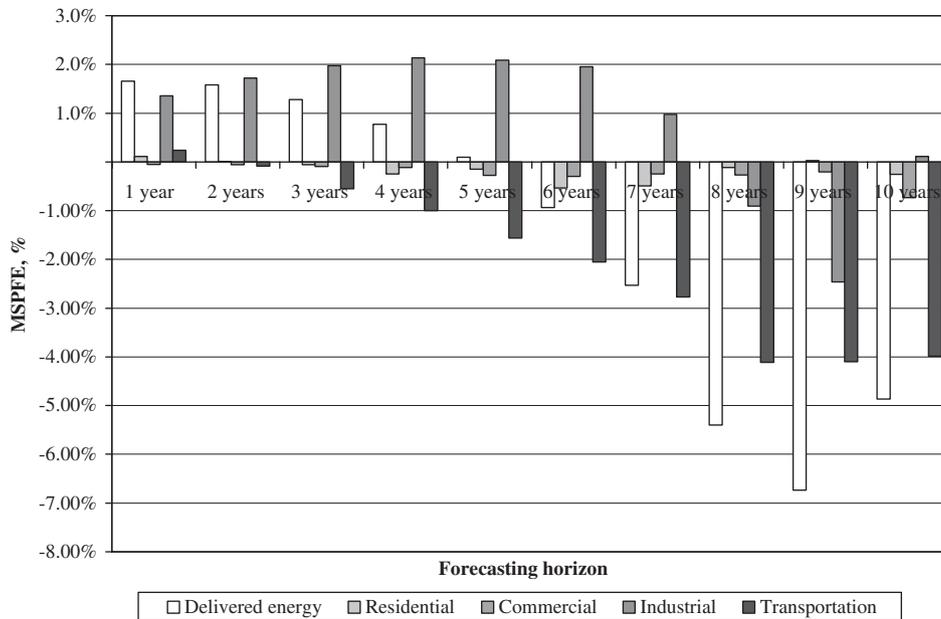


Fig. 3. MSFPE for energy consumption by forecast length.

4.3. Analysis of MSFPE by sector

We apply the MSFPE discussed earlier to help partition out the sectoral contributions of each sector on the total forecast error. Such analysis identifies sectors that have “high leverage” with respect to total error.

Table 1 shows MSFPE for each sector. Recall that MSFPE demonstrates a sector’s average, real contribu-

tion to the total energy forecast error for a given forecast length. As shown in Fig. 3, the transportation sector is a major contributor to total energy forecast error, particularly for forecasts with longer time horizons. In some cases, transportation underestimation by itself exceeds the total forecast underestimation, since the error is reduced by overestimation in other sectors. For forecasts with time horizons less than five years, the industrial sector is the largest contributor to this error.

4.4. Analysis of forecast improvements

For a given forecast horizon, we can analyze error trends to determine if forecasts are improving over time. For example, we may ask: *Have 7-year forecasts improved in accuracy from 1982 to 1996?* Here, we conduct this analysis for several forecast time horizons (3-, 5-, and 7-year) across sectors.

Figs. 4 through 6 present the results of this analysis. Each graph shows the absolute error based on the year in which the forecast was made. So, for example, Fig. 4 evaluates whether three-year forecasts improved over time (from 1982 to 2000). Similarly, using Figs. 5 and 6,

we can determine whether five-year and seven-year forecasts have improved.

The figures show that there is a general randomness in this error; but it is interesting to note that all graphs are bowl shaped with minimum errors around AEO 1987 – AEO 1989. The reasons behind this concave pattern need to be explored further.

5. Conclusion

This paper evaluates forecast error for US energy forecasts made between 1982 and 2003. Forecast

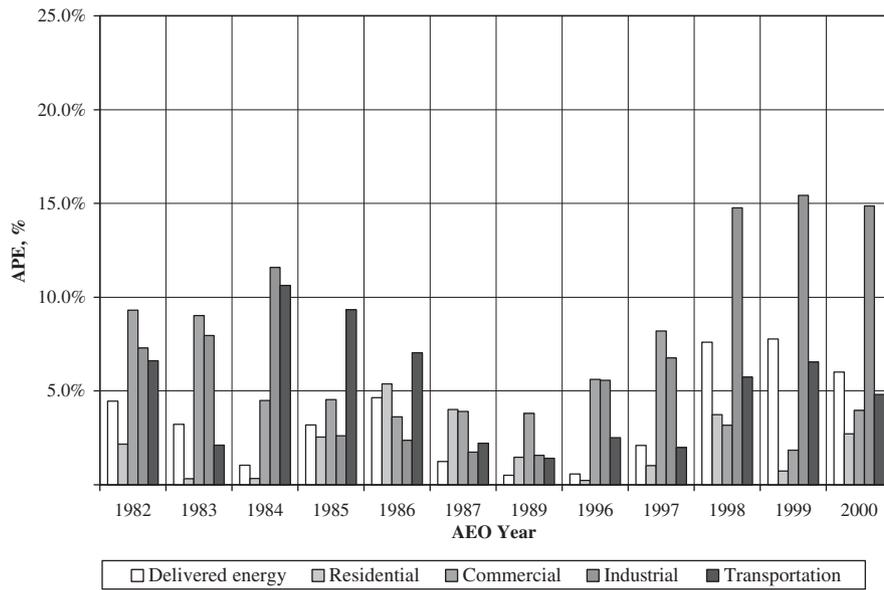


Fig. 4. Absolute errors in energy consumption by sector for 3-year forecasts by year of forecast.

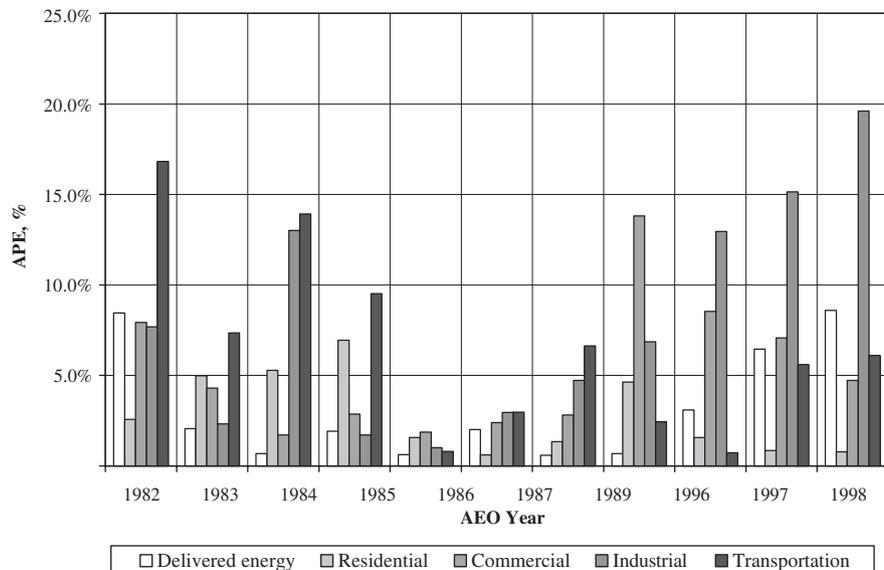


Fig. 5. Absolute errors in energy consumption by sector for 5-year forecasts by year of forecast.

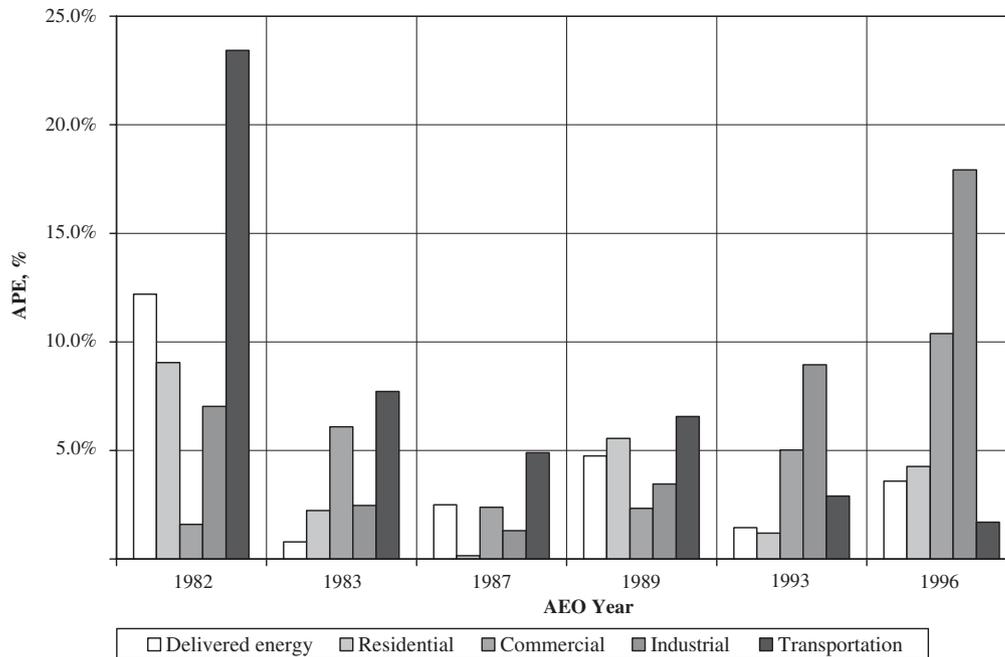


Fig. 6. Absolute errors in energy consumption by sector for 7-year forecasts by year of forecast.

time horizons ranged from one to ten years. Based on our analysis, we can draw the following conclusions.

- Using total (aggregate) energy consumption forecast errors to judge the quality of US energy forecasts is misleading. In fact, these relatively low, aggregate forecast errors conceal much higher errors at the sector (disaggregate) level. For example for the 5-year forecasts made between 1982 and 1998, the MPE for total energy consumption was 0.1%. Yet, this hides the fact that the industrial sector was on average overestimated by 5.9% and the transportation sector was underestimated by 4.5%.
- The residential sector errors are the lowest among all sectors. Commercial sector errors are higher than residential sector but only influence total consumption errors slightly (as shown by the low MSPFE). Meanwhile, the industrial and transportation sector errors are the highest and offer the largest contribution to total error.
- Both the commercial and transportation sector were consistently underestimated.
- We found no evidence that energy forecasts for the time period studied are becoming more accurate over time.
- Our analysis showed systemic errors in all sectors, but we do not attempt to explain the reasons of such behavior. It may be a result of incorrect core assumptions, incorrect relationships between model parameters, or major shifts in economic, politic or demographic conditions.

Forecasting is tricky business. This is particularly true in the energy field, where the highly random behavior of energy prices and technological change make forecasting difficult. However, because these forecasts are so integral to policy and business decisions, it is worth analyzing where these forecasts fail. Our results indicate that certain energy sectors (namely the industrial and transportation sectors) seem to exhibit systemic modeling problems that should be further explored. Future work in this area may include a more disaggregate analysis that evaluates the assumptions and inputs for all sectors and especially for the transportation and industrial sector models. These two sectors suffer from the largest errors and tend to influence total energy consumption error the most. Further research may help answer whether errors in these sectors are driven by errors in inputs, model equations, or both.

References

- Ascher, W., 1978. *Forecasting: An Appraisal for Policy-Makers and Planners*. Johns Hopkins University Press, Baltimore MD.
- Craig, P., Gadgil, A., Koomey, J., 2002. What can history teach us? A Retrospective Examination of Long-Term Energy Forecasts for the United States. *The Annual Review of Energy and Environment* 27, 83–118.
- Energy Information Administration, 1982–2003. *Annual Energy Outlook*. DOE/EIA-0383. US Department of Energy, Washington, DC.
- Energy Information Administration, 1998. *Issues in Midterm Analysis and Forecasting*. DOE/EIA-0607(98). US Department of Energy, Washington, DC.
- Energy Information Administration, 2003a. *Annual Energy Review*. DOE/EIA-0384. US Department of Energy, Washington, DC.

- Energy Information Administration, 2003b. The National Energy Modeling System: An Overview. Accessed on November 3 at <http://www.eia.doe.gov/oiaf/aeo/overview/overview.html>.
- Energy Information Administration, 2004. Annual Energy Outlook 2004 with Projections to 2025. DOE/EIA-0383. US Department of Energy, Washington, DC.
- Holte, S. H., 2002. Annual energy Outlook Forecast Evaluation. Accessed on November 28 at http://www.eia.doe.gov/oiaf/archive/analysispaper01/forecast_eval.html.
- Linderoth, H., 2002. Forecast errors in IEA-countries' energy consumption. *Energy Policy* 30, 53–61.
- Munson, R., 2004. Improving prediction of energy futures. *Issues in Science and Technology* XX, 26–29.
- O'Neill, B., Desai, M., 2005. Accuracy of past projections of US energy consumption. *Energy Policy* 33, 979–993.
- Sanchez, E., 2002. Annual Energy Outlook Forecast Evaluation. Accessed on November 28 at http://www.eia.doe.gov/oiaf/analysispaper/forecast_eval.html.
- Smil, V., 2000. Perils of Long Range Energy Forecasting: Reflections on Looking Far Ahead. *Technological Forecasting and Social Change* 65, 251–264.
- Smil, V., 2003. *Energy at the Crossroads*. MIT Press, Cambridge, MA.