Density Forecast Estimation for Long-term Energy Market Projections

Janice Lent[1], Adebowale J. Sijuwade[1], Caitlin Steiner[1]
Sauleh A. Siddiqui[1]
[1]U.S. Energy Information Administration,[1]
1000 Independence Ave., S.W., Washington DC 20858

Abstract
We discuss methods of estimating density forecasts and prediction intervals for long-term energy market projections (e.g., energy prices, production, and consumption) published in the U.S. Energy Information Administration's (EIA's) *Annual Energy Outlook*. Following Kaack et al. (2017), we examine density forecasts based on previous projection errors and historical volatility, using both empirical distributions and assumed Gaussian distributions. Because of the small sample sizes of previous projection errors that are available for these projected series, we also consider empirically-based density forecasts with exponential smoothing and trend extrapolation applied to the initial error bounds. We use estimated regression coefficients, combined with metrics of historical volatility, to develop uncertainty metrics for projected series with no available data on previous projection errors.

Key Words: density forecast, energy market projection, time series, exponential smoothing

1. Background: EIA's *Annual Energy Outlook*

Each year, the U.S. Energy Information Administration (EIA) publishes long-term energy market projections in its *Annual Energy Outlook*. EIA analysts base the projections on output from EIA's National Energy Modeling System (NEMS), a modular computer-based system that simulates the energy supply, demand, and conversion processes of U.S. energy markets.

EIA uses the NEMS to project the effects that economic, environmental, and global energy factors may have on the U.S. energy system. The current NEMS projection period runs from the present year through 2050. Annual projections include energy production, consumption, and prices, for a variety of fuels. EIA analysts develop a variety of *cases*, each incorporating a different set of assumptions (e.g., high or low oil prices, high or low macroeconomic growth) and produce a different set of projections for each case. The AEO Reference case represents a basis of comparison for other cases, which we call *side cases*, and assumes current energy policies and moderate GDP growth. EIA also uses the NEMS to project the impact of new energy policies under consideration.

Figure 1 is a schematic representation of the NEMS. The system incorporates modules for energy supply (oil, natural gas, etc.), demand (residential, transportation, etc.) and conversion (electricity and liquid fuels). It also simulates the effects of global energy markets and macroeconomic conditions on the U.S. energy system. EIA's website offers extensive NEMS documentation [1].

Each year's AEO presents projections for the Reference case and several side cases, as well as extensive analysis and discussion of current and expected energy market conditions and the effects

[1] The analysis and conclusions expressed in this paper are those of the authors and do not represent the official position of the U.S. Energy Information Administration (EIA) or the U. S. Department of Energy (DOE). Caitlin Steiner conducted this research when she was a Mathematical Statistician at the U.S. Energy Information Administration, though she currently is employed by the U.S. Department of State.

of any recent energy policy changes. EIA analysts vary the side cases and input assumptions each year. Typical AEO side cases include high and low cases for oil prices, macroeconomic growth, and technological advancement. The AEO assumptions, along with data limitations and other factors, are sources of uncertainty in AEO projections.

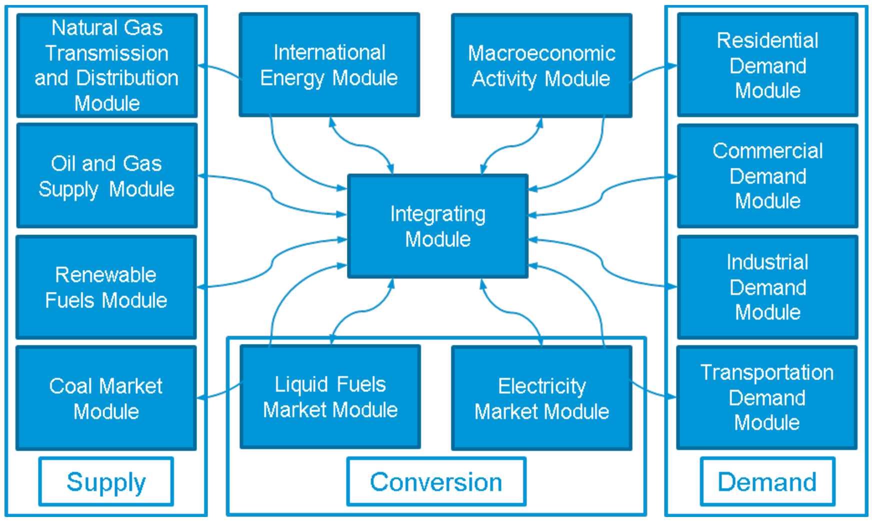


Figure 1: Schematic representation of EIA's National Energy Modeling System (NEMS)

2. Motivation: Quantifying Uncertainty in AEO Projections

2.1 Density Forecasts for Long-term Projections

Users of long-term projections increasingly rely on density forecasts, or uncertainty metrics, to evaluate point projections and determine their appropriate uses. These metrics often indicate the level of uncertainty based on how far off the analysts' projections have been in the past, on average, for each projection horizon. The uncertainty bounds are interpreted as probabilistic confidence intervals; we estimate, for example, a 95% probability that the actual values will fall between the 95% uncertainty bounds.

Methods of quantifying uncertainty in long-term energy projections, through density forecasting, are discussed in the economics literature. Shlyakhter et al (1994) develop a one-parameter model to estimate a probability distribution for projections. Sherwin et al. (2019) conjecture that the U.S. energy system has become harder to predict and more volatile in the past decade, in comparison to the past two to three decades. In the AEO Reference case projections, a large percentage of the highest projection errors within the past few decades occur within the most recent decade. They also note the high frequency of large year-to-year changes in key energy quantities. Shlyakhter et al. (1994) present a density forecasting method based on a combination of past forecasting errors and either (a) expert judgment of uncertainty bounds or (b) projections from other AEO cases (e.g., the high and low oil price cases).

Kaack et al. (2017) examine four main density forecasting methods that are independent of expert judgment (either explicit or in the form of AEO case projections) and therefore replicable:

1. Nonparametric Empirical Prediction Interval (denoted NP_1)
2. Nonparametric Centered Empirical Prediction Interval (denoted NP_2), centered on the current projected series
3. Gaussian Distribution, centered on the current projected series, with standard deviations of the past projection errors (denoted G_1)

4. Gaussian Distribution, centered on the current projected series, with standard deviations of the past historical proportional changes over time (denoted G_2)

Appendix B gives the formulas for methods 2 through 4, which we discuss in more detail below. To compare the density forecasts, Kaack et al. (2017) use the continuous ranked probability score (CRPS), also discussed in Appendix B. Gneiting and Raftery (2007) discuss several alternative scoring measures similar to the CRPS.

Kaack et al. (2017) define the projection horizon H_i to be the number of years between the AEO publication year and the projected year, plus 1 (e.g., for AEO 2020, the H_0 projections are for 2019). They present analysis for 18 projected AEO series and for projection horizons $H = 2$ to $H = 9$. They find that methods G_1 and G_2 perform best overall of the four methods considered, as assessed by CRPS scores. Of the nonparametric distribution methods, NP_2 performed better than NP_1.

In our analyses, we focus on the utility of the different methods rather than on their CRPS scores. Cones of uncertainty computed by method NP_1, which is based on quantiles of the relative projection errors from past years and not centered on the current projected series, may, in some cases, fail to encompass the full current projected series. Because such cones would be unintuitive for AEO readers, we exclude method NP_1 from our evaluation. Method G_2, while based on the implicit assumption that increased historical volatility decreases projection accuracy, is of particular interest from a practical standpoint, because the G_2 uncertainty metrics can be computed from historical data, without the need for AEO Retrospective analysis, which is resource intensive. Method G_1 is simple to apply, given the Retrospective analysis. It assumes that the relative projection errors follow a Gaussian distribution, an assumption that appears justified (see Appendix A). Method NP_2 requires no distributional assumption and indicates the direction, in addition to the magnitude, of past projection errors.

2.2. Implementing Density Forecasts in the AEO

We present research on three density forecasting methods, as applied to the AEO Reference case projections:

1. Method NP_2, based on quantiles of the centered empirical distributions of the relative projection errors
2. Method G_1, based on Gaussian quantiles multiplied by the standard deviations of the relative projection errors
3. Method G_2, based on Gaussian quantiles multiplied by the standard deviations of the relative changes over time in historical values

Subsection 2.3 provides graphical examples of uncertainty metrics computed by the above methods. Our goal is to develop practical methods of implementing the uncertainty metrics in the AEO. Because, two of the three methods rely on AEO retrospective analysis, we face two main challenges to implementation, based on the limited amount of historical data published in the AEO Retrospective:

First, for each series analyzed in the most recent AEO Retrospective (from 2022), only a limited number of AEO projection-to-actual data comparisons are available. The small number of data points available for high projection horizons, together with their autocorrelation, causes the uncertainty cones to become unstable and often collapse. As discussed in Section 3, we apply time-series smoothing and prediction techniques to the uncertainty bounds to stabilize them and prevent them from collapsing at high projection horizons.

Second, the AEO 2022 Retrospective includes comparisons for 31 AEO projected Reference case series, but many more Reference case series are projected in each year's AEO. For AEO projected series not represented in the Retrospective analysis, we compute uncertainty metrics based on a Gaussian error assumption and historical volatility, using Kaack's method G_2. The underlying assumption is that the more volatile series are more difficult to project. Our research indicates that the G_2 metrics, in most cases, tend to overstate uncertainty relative to Kaack's method G_1, which is based on a Gaussian error assumption applied to the Retrospective comparisons. As discussed in Section 4, we use regression analysis to estimate adjustment factors that we can apply to the G_2 uncertainty cones to approximate G_1 cones. Because we compute the G_2 cones from historical data only, we can compute them for AEO projected series not included in the AEO Retrospective.

2.3 Examples of Density Forecasts Based on Retrospective Analysis

To illustrate the methods described in subsection 2.2, we present graphs computed by methods NP_2 and G_1. Appendix B gives the technical details of the calculations.

2.3.1 Energy Consumption Series

We examined the uncertainty metrics for projections of residential, commercial, industrial, and transportation consumption. Figure 2.3.1 shows results for the residential sector. Due to the small sample sizes and autocorrelated data points, the cones tend to collapse for the higher horizons, motivating the smoothing procedure we introduce in subsection 3.1 below.

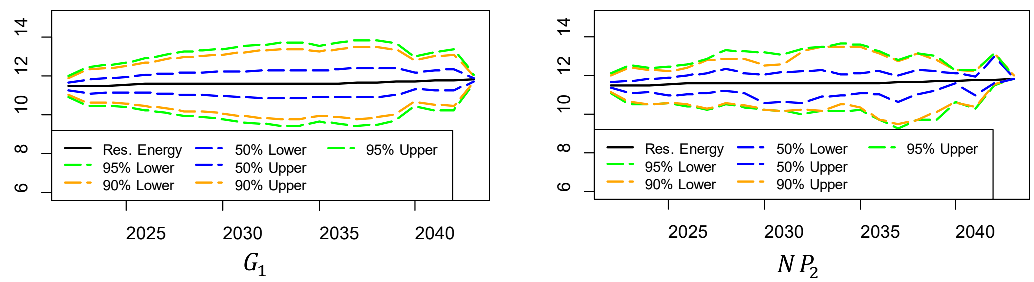


Figure 2.3.1: Uncertainty cones for residential energy consumption (quadrillion Btu), methods G_1 and NP_2.

While the G_1 cones are symmetric, based on Gaussian quantiles and the standard deviations of the relative projection errors, the NP_2 cones are based on empirical quantiles of the relative projection errors and can therefore indicate the direction, as well as the magnitude of the projection errors. In general, the G_1 cones appear to be conservative relative to the NP_2 cones, i.e., they indicate somewhat larger projection errors.

2.3.2 Energy Price Series

Figure 2.3.2 shows the results of methods G_1 and NP_2 for projections of natural gas prices to the electric power sector. Because we assume prices follow an approximately lognormal distribution, we analyze the price data in log scale, and then transform back, which gives the cones a somewhat different shape.

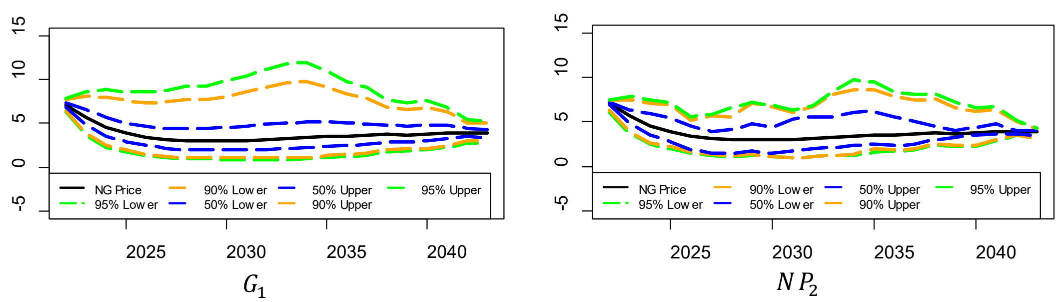
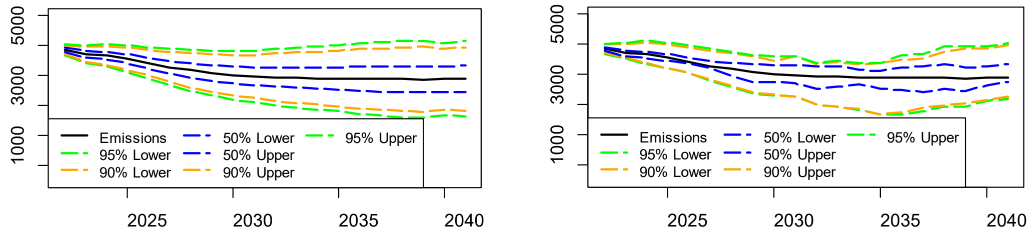


Figure 2.3.2: Uncertainty cones for natural gas prices to the electric power sector (dollars per million Btu), methods G_1 and NP_2.

Again in this case, the Gaussian (G_1) cones are conservative relative to the empirical (NP_2) cones. Because the prices are analyzed in log scale, the general direction of the errors is difficult to discern from the NP_2 cones. Again we see the cones computed by both methods collapsing at higher projection horizons.

2.3.3 Macroeconomic Series

The large number of series projected for the AEO includes gross domestic product (GDP), energy intensity (energy consumption per unit of GDP), and energy-related carbon dioxide (CO_2) emissions. Figure 2.3.3 shows the uncertainty cones for energy-related CO_2 emissions, computed by methods G_1 and NP_2. The NP_2 cones clearly show that emissions have been over-projected, rather than under-projected in previous years.



G_1 NP_2

Figure 2.3.3: Uncertainty cones for energy-related CO_2 emissions (millions of metric tons), methods G_1 and NP_2.

3. Estimation Methods

3.1 Smoothing Uncertainty Bounds

As seen in the graphs above, the uncertainty bounds for the later horizons are often unreliable, primarily due to small sample sizes and autocorrelated observations. We explored methods of smoothing the empirical and Gaussian uncertainty bounds for horizons 5 through approximately 10 and trend-projecting the smoothed series out to horizon 15 to 20, where we adopt Kaack's terminology for the horizons. For some series with particularly unstable cones, we reduced the length of the smoothed series (below horizon 10) and increased the length of the projected series.

3.1.1. Smoothing Method

To smooth the uncertainty cones, we used the R `HoltWinters` function, which implements an exponentially weighted filtering of the level, trend, and seasonal components of a time series. It generally involves three parameters labeled *alpha*, *beta*, and *gamma,* which take values in the half-open interval $(0, 1]$. The *alpha* parameter controls the exponential smoothing of the series, while the *beta* parameter controls the smoothing of the estimated trend component. We set the *gamma* parameter, which controls the smoothing of the estimated seasonal component, to `FALSE` (implementing Holt smoothing), because our application involves only annual data. If users specify no values for the *alpha* and *beta* parameters, the R `HoltWinters` function computes parameters that minimize the squared one-step-ahead prediction errors for the series.

To smooth each uncertainty bound Y, we first identify two break points Y_a to Y_b, that divide the series into three sections:

1. The *original* section runs from H_0 to the first break point Y_a. We leave this section unchanged, because the sample sizes at low horizons are generally sufficient to provide good results.
2. The *smoothed* section runs from the first break point Y_a to the second break point Y_b. We use the `HoltWinters` function to apply exponential smoothing to this section.
3. The *predicted* section runs from the second break point Y_b to the end of the projected series. For the second break point Y_b, we choose the point at which the uncertainty cone begins to collapse. Rather than allowing the cone to collapse, we use the `predict` function in R to

extend the estimated trend from the smoothed section. We use a weighted average of the original and predicted series for this section of the uncertainty bound.

Although the break points Y_a and Y_b vary by series, we automated the above process for Gaussian uncertainty bounds.

3.1.2 Smoothing Example

Figure 3.1.2 shows the NP_2 uncertainty cones for energy-related CO_2 emissions, before and after smoothing. In this case the break points (Y_a and Y_b) were approximately 2027 and 2035. In particular, the lower uncertainty bounds begin to collapse at around 2035 in the unsmoothed series. The trend projection prevents the cones from collapsing for the later years of the projection period. Because some data are available for these years, however, we used a weighted average of the trend projection and the original series for the years beyond 2035.

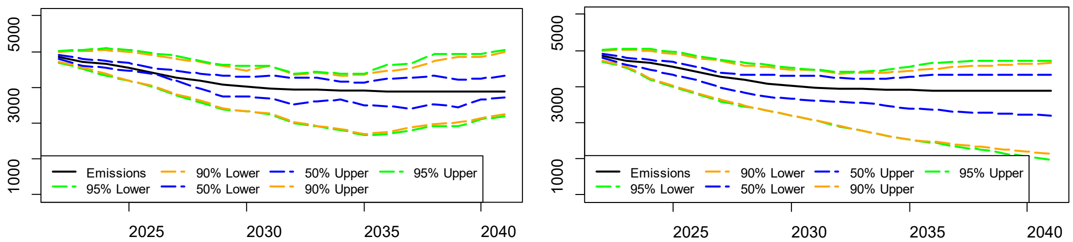


Figure 3.1.2: NP_2 uncertainty cones for energy-related CO_2 emissions (millions of metric tons), before and after smoothing

3.2 Using Historical Volatility to Approximate Gaussian (G_1) Cones

EIA performs AEO Retrospective analysis every two to three years and includes a limited number of Reference case series (e.g., 31 series for the AEO 2022 Retrospective). We can therefore estimate uncertainty cones by methods NP_2 and G_1 only for a limited number of projected AEO series. Method G_2, however, relies only on measures of volatility computed from historical data, so we can estimate the G_2 cones for all series projected in the AEO. In this section, we present a method of estimating G_1 cones by adjusting the standard deviations computed for the G_2 cones.

Figure 3.2.1 shows uncertainty cones for industrial energy consumption, estimated by method G_2, (historical volatility) and Figure 3.3.2 (a) shows the corresponding cones computed by method G_1 (retrospective analysis). The G_2 cones are clearly wider than the G_1 cones. To adjust for the extra width, we compute regression coefficients, based on the retrospective and historical data for *residential* energy consumption. For the residential energy data, we fit Model 1:

$$\sigma_{1,y} = \beta_0 + \beta_1\sigma_{2,y} + \varepsilon, \qquad (3.2.1)$$

where, for the variable y, $\sigma_{1,y}$ is the standard deviation of the projection errors (for the G_1 cones), $\sigma_{2,y}$ is the standard deviation of the historical volatility (for the G_2 cones), and ε is a normally distributed error term. Because we intuitively conclude that more volatile energy-related series are more difficult to project than stable series, Model 1 assumes a positive correlation between the historical volatility of an energy-related series ($\sigma_{2,y}$) and the variability of the resulting projection errors ($\sigma_{1,y}$).

We then use the estimated coefficients $\hat{\beta}_0$ and $\hat{\beta}_1$ to estimate the G_1 standard deviations for industrial energy consumption. We compute

$$\hat{\sigma}_{1,i} = \hat{\beta}_0 + \hat{\beta}_1\sigma_{2,i}, \qquad (3.2.2)$$

where $\sigma_{2,i}$ is the vector of standard deviations computed for industrial energy consumption (pooled across horizons) by method G_2, and $\hat{\sigma}_{1,i}$ is the estimated (based on residential energy consumption retrospective data) vector of G_1 standard deviations for industrial energy consumption. Figure 3.2.2 (b) shows the adjusted uncertainty cones, which are similar in width to the G_1 cones computed for industrial energy consumption (Figure 3.2.2 (a)).

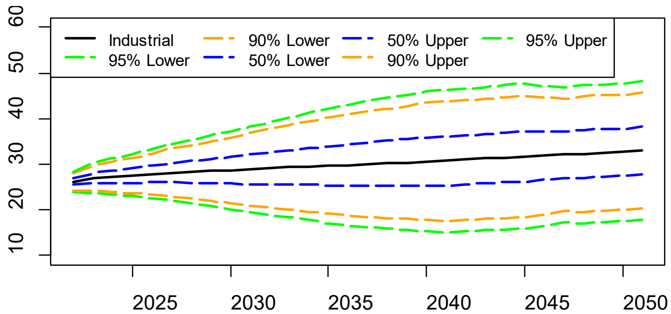
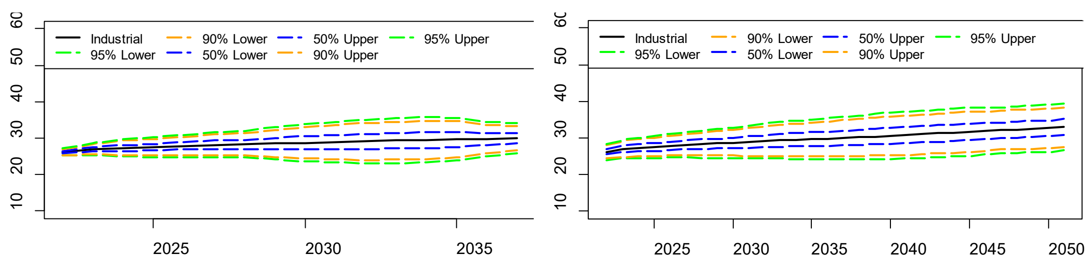


Figure 3.2.1: Gaussian uncertainty cones for industrial energy consumption (quadrillion Btu), estimated based on historical volatility (method G_2)



(a) G_1 (b) Adjusted G_2

Figure 3.2.2: Gaussian uncertainty cones for industrial energy consumption (quadrillion Btu), (a) estimated based on Retrospective analysis (method G_1); and (b) estimated based on historical volatility (method G_2) and adjusted using Residential consumption coefficients from Model 1

Figure 3.2.3 shows the standard deviations, by horizon, for commercial energy consumption. The blue (middle) line shows the standard deviations of the Retrospective projection errors (method G_1), while the grey (top) line shows the standard deviations of the historical changes (method G_2). To obtain estimated G_1 standard deviations (the orange line), we adjusted the G_2 standard deviations using the model coefficients from Model 1, fit on data for the residential sector. The adjustment successfully brings the standard deviations close to the correct level for method G_1 (the blue line).

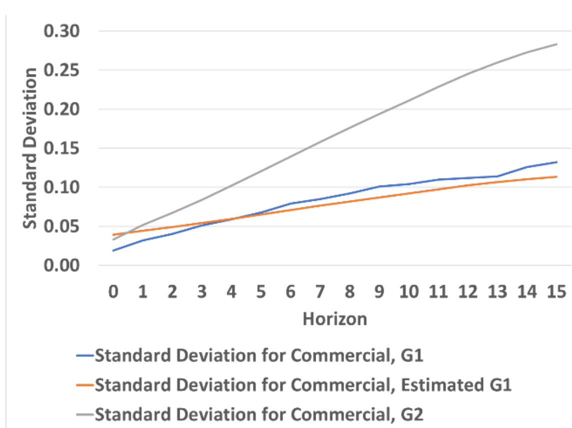


Figure 3.2.3: Standard deviations for Commercial energy consumption, estimated from standard deviations for Residential energy consumption, by horizon

To generalize the method of estimating G_1 standard deviations by adjusting G_2 standard deviations, we developed a collection of robust adjustment factors for categories of AEO Reference case projection series, to adjust G_2 cones to approximate G_1 cones.

1. We computed G_1 and G_2 standard deviations for all series in the 2022 AEO Retrospective.
2. We tabulated regression coefficients, p-values, and adjusted R^2 statistics for Model 1:

$$\sigma_{1,y} = \beta_0 + \beta_1 \sigma_{2,y} + \varepsilon, \tag{3.2.3}$$

where, for the variable y, $\sigma_{1,y}$ is the standard deviation of the projection errors (for the G_1 cone), $\sigma_{2,y}$ is the standard deviation of the historical volatility (for the G_2 cone), and ε is a normally distributed error term. For each series, we pool the standard deviations across horizons to fit the regression models. Table A1 (Appendix A) gives estimated coefficients and diagnostics for Model 1.

3. We also tabulated regression coefficients, p-values, and adjusted R^2 statistics for Model 2 (regression through the origin):

$$\sigma_{1,y} = \gamma \sigma_{2,y} + \varepsilon, \tag{3.2.4}$$

where the notation is as for Model 1. The coefficient γ in Model 2 indicates whether the adjustment should make the G_2 cone narrower or wider. Table A2 (Appendix A) gives estimated coefficients and diagnostics for Model 2.

4. We examined the table of summary diagnostics and generalize to obtain robust, somewhat conservative, adjustment factors based on the 2022 Retrospective data.

5. We used the robust adjustment factors, along with historical data, to estimate G_1 cones for most of the series in Tables 1 through 11 of AEO 2023. The AEO tables are available on the EIA website.

4. Results

4.1 Smoothing Results

Here we provide more examples of uncertainty bounds that have been smoothed by the method detailed in subsection 3.1.1. Figure 4.1.1 shows the NP_2 uncertainty bounds for delivered residential energy consumption. While the unsmoothed bounds become unstable and collapse toward the end of the series, the smoothed bounds expand.

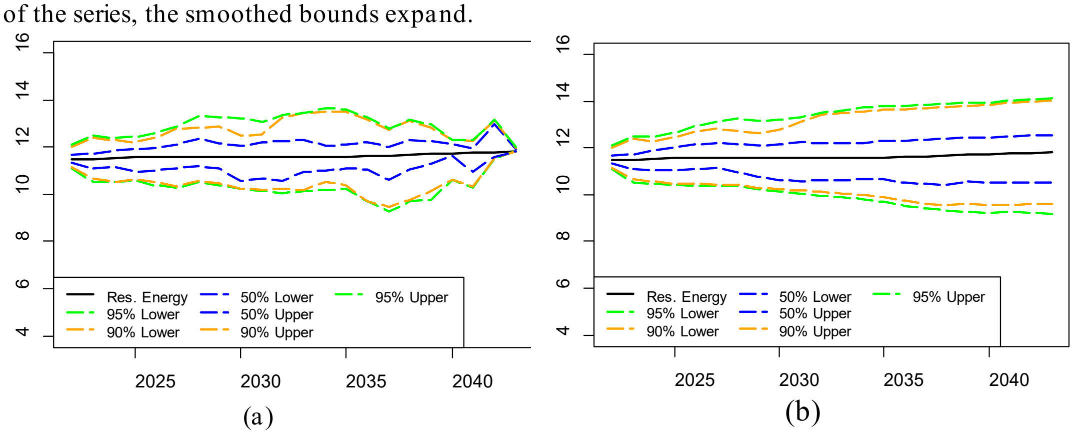
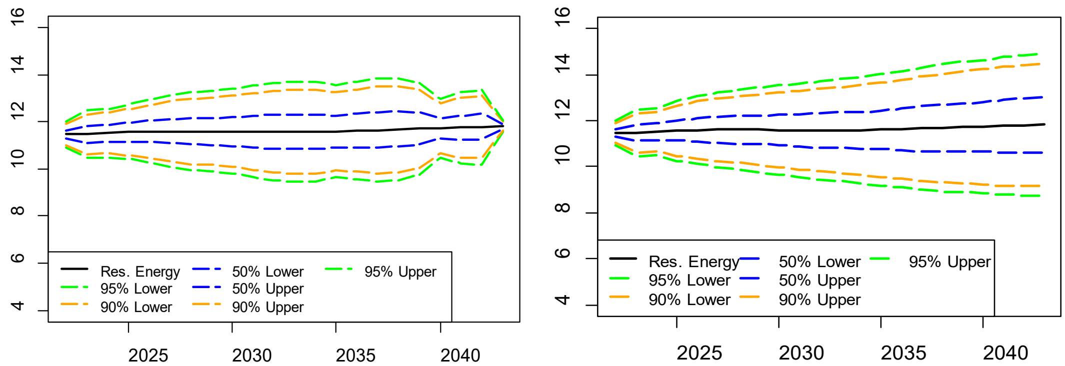


Figure 4.1.1: NP_2 uncertainty cones for residential energy consumption (quadrillion Btu), (a) unsmoothed and (b) smoothed

Figure 4.1.2 shows the original and smoothed G_1 (Gaussian) uncertainty bounds for delivered residential energy consumption. Comparing Figure 4.1.1(b) with Figure 4.1.2(b), we find that the smoothed G_1 uncertainty bounds are similar to the smoothed NP_2 uncertainty bounds. For most series in the 2022 Retrospective, the G_1 cones provide a reasonable approximation of the NP_2 cones.

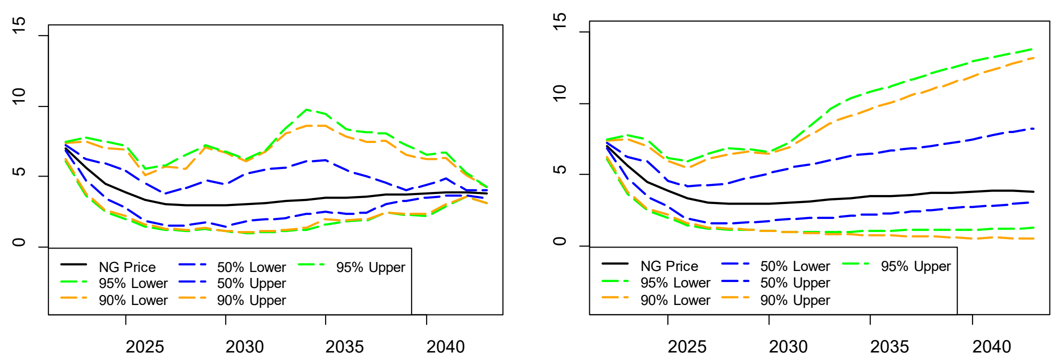


(a) (b)

Figure 4.1.2: G_1 (Gaussian) uncertainty cones for projected residential energy consumption (quadrillion Btu), (a) unsmoothed and (b) smoothed

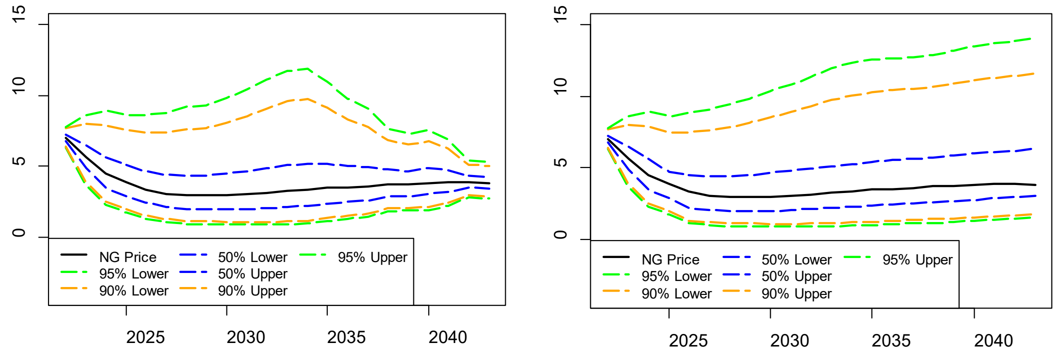
Because we assume that prices follow a lognormal distribution, we analyze the price series in log scale. Figures 4.1.3 and 4.1.4 show the original and smoothed NP_2 and G_1 uncertainty cones for nominal natural gas prices to the electric power sector (dollars per million Btu). Again, we note the similarity between the smoothed NP_2 and G_1 cones.

For AEO projected series not included in the Retrospective analysis, we can only compute the estimated G_1 cones, as discussed in subsections 3.2 and 4.2 (below). After smoothing, however, the Gaussian cones provide a reasonable approximation of the empirical NP_2 cones for most series.



(a) (b)

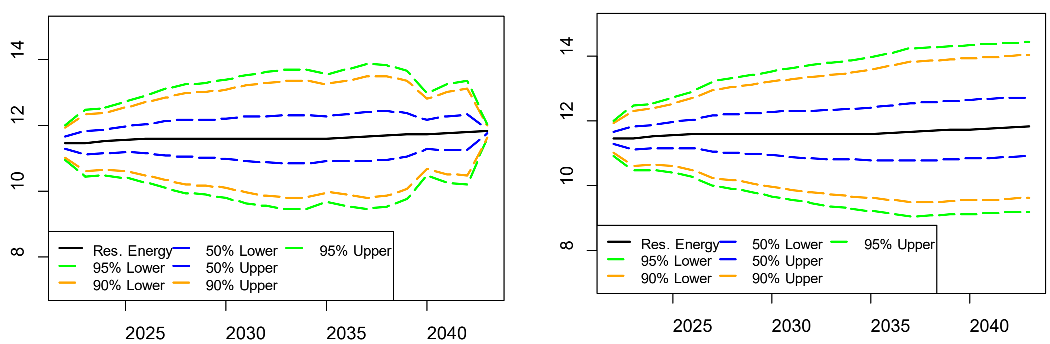
Figure 4.1.3: NP_2 uncertainty cones for projected nominal natural gas prices to the electric power sector (dollars per million Btu), (a) unsmoothed and (b) smoothed. Prices are analyzed in log scale.



(a) (b)

Figure 4.1.4: G_1 uncertainty cones for projected nominal natural gas prices to the electric power sector (dollars per million Btu), (a) unsmoothed and (b) smoothed.

For the Gaussian uncertainty cones, we automated the smoothing algorithm. The automated algorithm smooths and projects the series of standard deviations, ensuring a monotone increasing series for the higher horizons. Based on default function parameters, the algorithm searches in specific ranges for appropriate break points between the three sections. For most series, default function parameters work well. We use an alternative set of function parameters for some series. Once the algorithm finds the appropriate break points, the smoothing is relatively straightforward. Figure 4.1.5 shows the G_1 uncertainty cones for residential energy consumption, (a) unsmoothed and (b) smoothed by the automated algorithm.



(a) (b)

Figure 4.1.5: G_1 uncertainty cones for projected residential energy consumption (quadrillion Btu), (a) unsmoothed and (b) smoothed by the automated algorithm.

4.2 Generalizing Regression Coefficients to Estimate G_1 Uncertainty Cones

For those AEO series for which we have no retrospective data, we can compute uncertainty metrics by method G_2, based on historical data. We used the data in the 2022 retrospective to develop simple adjustment factors that we can apply to the G_2 standard deviations to approximate G_1 standard deviations for those series without retrospective analyses. The factors differ by type of series (e.g., consumption, price).

4.2.1. Regression Results

Table 1 shows the γ coefficients from Model 2 for the 31 projected series included in the 2022 AEO Retrospective. The coefficients indicate the accuracy of the AEO projections (G_1 standard deviations), relative to the volatility of the historical series (G_2 standard deviations). These can be used to adjust G_2 standard deviations (based on historical data) to approximate G_1 standard deviations. The table gives an approximate rating of projection uncertainty for each series, relative to historical volatility:

1. High, due to large changes not projected in the Reference case ($\gamma > 1$)
2. Medium, i.e., consistent with volatility (γ close to 1)
3. Slightly low relative to volatility ($0.7 < \gamma < 0.9$)
4. Moderately low relative to volatility ($0.4 < \gamma < 0.7$)
5. Very low relative to volatility ($0.1 < \gamma < 0.4$)
6. Extremely low relative to volatility ($\gamma < 0.1$)

Table 1: Estimated Gamma Coefficients (Model 2) and Categories of 2022 AEO Retrospective Series (EP = Electric Power sector)

	$\hat{\gamma}$	**Category**
High		
Petroleum net imports	8.2756	imports
Coal production excluding waste coal	2.0540	production
Total coal consumption	1.5677	consumption
Medium		
Energy intensity	1.1948	macro
Total energy-related CO_2 emissions	0.9934	macro
Slightly low		
Coal price to EP (nominal $)	0.8332	price
Imported cost of crude oil (nominal $)	0.7697	price
Imported cost of crude oil (constant $)	0.7094	price
Moderately low		
Coal prices to EP (constant $)	0.6572	price
Natural gas price to EP (nominal $)	0.6525	price
Natural gas price to EP (constant $)	0.5753	price
Crude oil production	0.5552	production
Transportation energy consumption	0.5417	consumption
Commercial energy consumption	0.4908	consumption
Industrial energy consumption	0.4887	consumption
Coal net generation (all sectors)	0.4776	generation
Residential energy consumption	0.4700	consumption
Dry natural gas production	0.4581	production
Average electricity prices (nominal $)	0.4360	price
Very low		
Petroleum and liquids consumption	0.3837	consumption
Natural gas net imports	0.3784	imports
Total energy consumption (all sectors)	0.3753	consumption
Total natural gas consumption	0.3378	consumption
Average electricity prices (constant $)	0.2492	price
Hydroelectric net generation	0.2214	generation

Total electricity sales excl. direct use	0.1780	consumption
Natural gas net generation (all sectors)	0.1370	generation
Extremely low		
Real GDP (cumulative growth)	0.0016	macro
Solar net generation (all sectors)	0.0015	generation
Wind net generation (all sectors)	9.04E-05	generation
Nuclear net generation (all sectors)	4.80E-05	generation

As shown in Tables A1 and A2 (Appendix A) the estimated $\hat{\beta}_1$ coefficients for Model 1 are similar in overall magnitude to the estimated $\hat{\gamma}$ coefficients from Model 2. For simplicity, we therefore use the $\hat{\gamma}$ coefficients to develop the robust adjustment factors, despite the significance of many of the intercept terms $\hat{\beta}_0$. We seek a one-step method of adjusting the G_2 standard deviations to approximate the G_1 standard deviations.

To generalize the $\hat{\gamma}$ values, we fit ordinary least squares regression models, using all 31 series, with $\hat{\gamma}$ as the dependent variable and various functions of the G_2 standard deviations as independent variables. We tried for example, the median standard deviation, the fourth moment of the relative differences (variance of the variance), and the G_2 sample size. None of the independent variables proved significant. Most had p-values exceeding 0.5, and the adjusted R^2 values for the models were less than 0.2. Table 1, however, shows a clear correlation between the $\hat{\gamma}$ coefficients and the series categories in the third column. Most of the consumption series, for example, fall in the "moderately low" and "very low" categories. A regression model with $\hat{\gamma}$ as the dependent variable and these categories as independent variables gives better results (detailed diagnostics are in Appendix A). Based on these results, we developed the simple adjustment factors shown in Table 2.

Table 2: Generalized Adjustment Factors for G_2 Standard Deviations, by Category

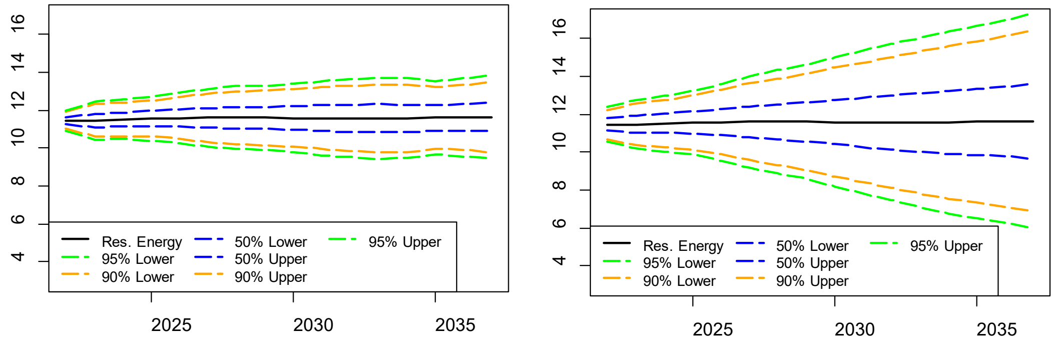
Category	Adjustment Factor
Price	0.7
Production or Consumption	0.5
Oil Imports	1
Other Imports	0.5
Electricity Generation	0.4
Macroeconomic	1

Because these factors are only approximate, they are conservative, deflating the G_2 standard deviations by a minimal amount to approximate G_1 standard deviations.

4.2.2. Estimating G_1 Cones from G_2 Cones

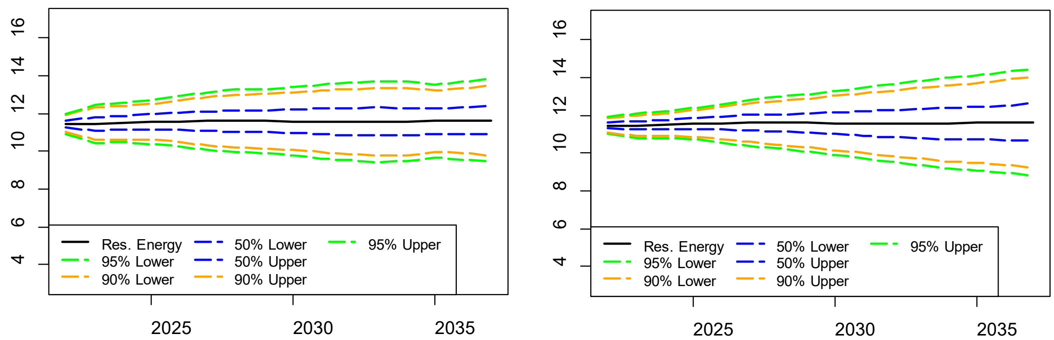
We apply the adjustment factors in Table 2 to estimate G_1 uncertainty cones for delivered residential energy consumption and natural gas prices to the electric power sector. Figure 4.2.2.1 shows the G_1 and unadjusted G_2 cones for residential energy consumption. The G_2 cones are clearly too wide to serve as estimates of the G_1 cones.

Table 2 gives an adjustment factor of 0.5 for consumption series. Applying this factor gives the uncertainty cones in Figure 4.2.2.1 (b), which approximate the G_1 cones in order of magnitude.



(a) (b)

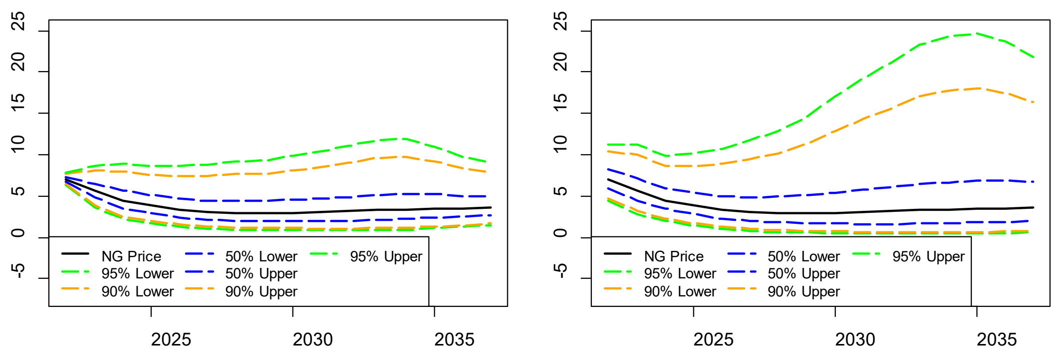
Figure 4.2.2.1: Uncertainty cones for projected residential energy consumption (quadrillion Btu) computed by (a) method G_1 (b) method G_2



(a) (b)

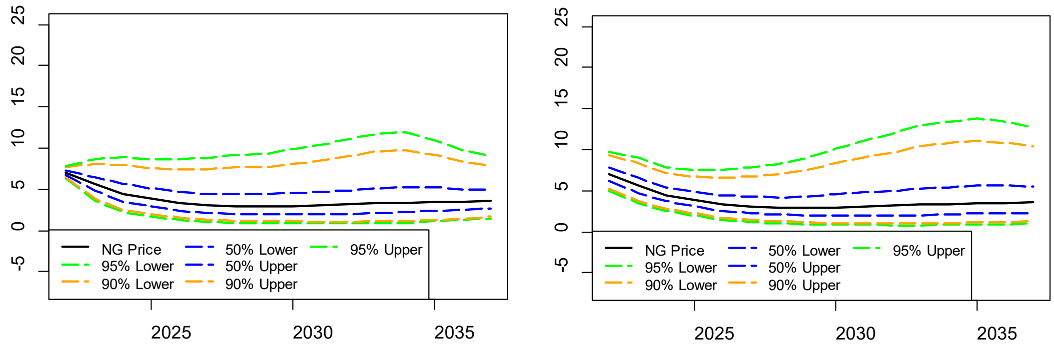
Figure 4.2.2.2: Uncertainty cones for projected residential energy consumption (quadrillion Btu) computed by (a) method G_1 (b) estimated method G_1

Similarly, Figure 4.2.2.3 shows the uncertainty cones for natural gas prices to the electric power sector, estimated by (a) method G_1 and (b) method G_2. Again, the G_2 cones are too wide. Table 2 gives an adjustment factor of 0.7 for price series. Applying this factor (with calculations in log scale) gives the uncertainty cones in Figure 4.2.2.4 (b), which approximate the G_1 cones in order of magnitude.



(a) (b)

Figure 4.2.2.3: Uncertainty cones for projected natural gas prices to the electric power sector (nominal dollars per million Btu) computed by (a) method G_1 (b) method G_2



(a) (b)

Figure 4.2.2.4: Uncertainty cones for projected natural gas prices to the electric power sector (dollars per million Btu) computed by (a) method G_1 (b) estimated method G_1

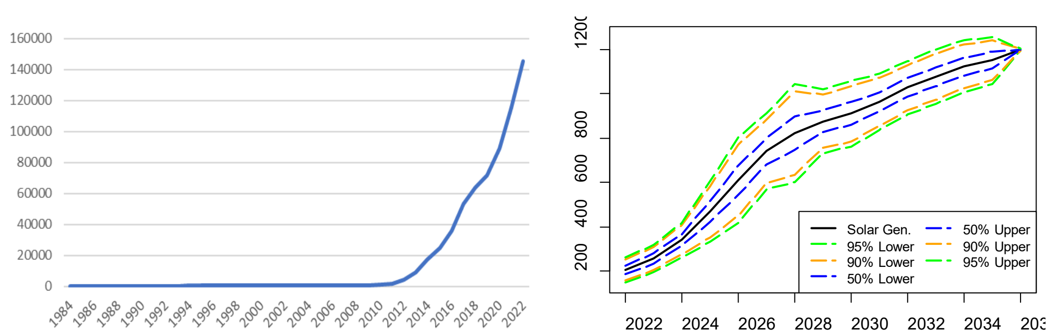
4.2.3. Series with Non-Gaussian Relative Historical Changes

The adjustment method illustrated above works well when both the relative projection errors used to compute the G_1 cones and the relative historical changes used to compute the G_2 cones follow Gaussian distributions, as assumed under the G_1 and G_2 methodologies. For all series in the AEO 2022 Retrospective, the relative projection errors are approximately Gaussian, as shown in Figure A.1 in Appendix A. For some series, however, the relative historical changes (inputs to the G_2 cones) fail the Shapiro-Wilk normality test, as shown in Figures A.2 through A.4 in Appendix A. We're currently developing special procedures to approximate the G_1 cones for these series.

The reasons behind the compromised Gaussian assumption vary by series. For example, Figure 4.2.3.1 shows the historical series for solar net generation (million Btu). Because of the rapid increase in solar generation from 2010 to the present, the historical relative changes for this series fail the normality test. Figure 4.2.3.2 shows the G_1 and G_2 uncertainty cones on the same scale. An extreme adjustment factor would be needed to adjust the G_2 cones for this series, and such factors are unlikely to be easily generalized to other series. Series with non-Gaussian relative historical changes make up 15% to 20% of the projected AEO Reference case series. The non-Gaussian series include:

- Solar, wind, and nuclear electricity generation
- Some other electricity generation series
- Some import and export series

As discussed in Section 5, we're currently testing alternative estimation procedures for these series.



(a) (b)

Figure 4.2.3.1: (a) Historical series for solar net generation (million Btu) and (b) G_1 uncertainty cones for projected solar net generation (billion kwh)

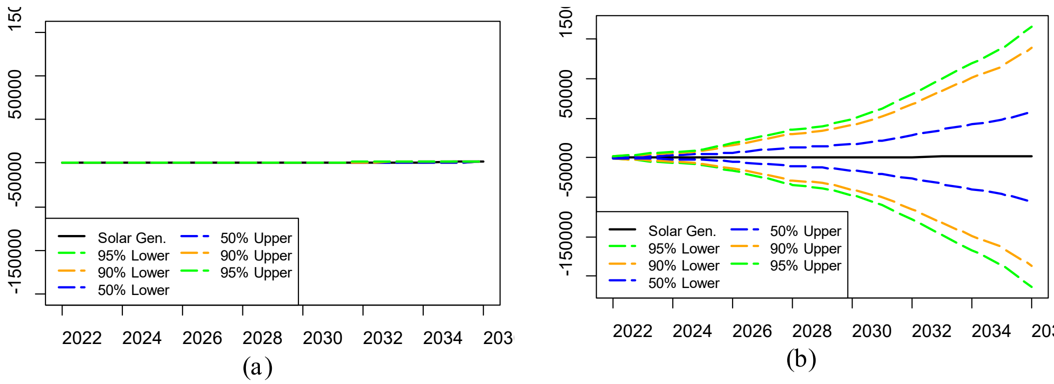


Figure 4.2.3.2: Uncertainty cones for solar net generation (billion kwh), on the same scale, for (a) method G_1 and (b) method G_2

5. Current Research

EIA's current research on AEO uncertainty measures focuses on two areas: (1) automating the smoothing and projecting algorithm for nonparametric (NP_2) uncertainty cones and (2) estimating G_1 uncertainty cones for series with non-Gaussian historical relative changes.

5.1 Automating the Smoothing and Projecting Algorithm for NP_2 Cones

Our automated algorithm for smoothing and projecting Gaussian uncertainty metrics smooths and projects the series of standard deviations of the relative projection errors. The algorithm ensures that the smoothed series of standard deviations is monotone increasing. We then compute the uncertainty cones based on the smoothed standard deviations and Gaussian quantiles.

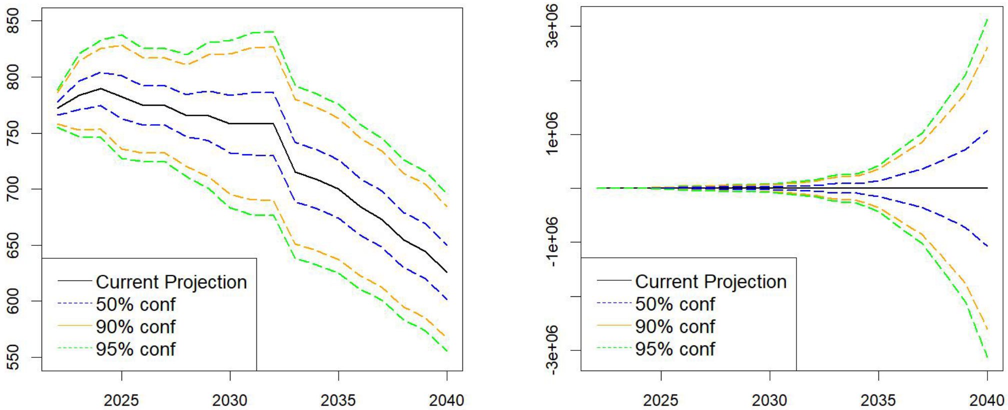
We can extend this method to the NP_2 cones by smoothing and projecting the series of differences between the NP_2 uncertainty bounds and the values of the projected series. Ensuring that the series of differences is monotone increasing will prevent the NP_2 cones from collapsing. We're currently working to automate the NP_2 smoothing algorithm.

5.2 Estimating G_1 Uncertainty Cones for Non-Gaussian Series

We're investigating several options for estimating G_1 uncertainty cones for series with non-Gaussian historical relative changes:

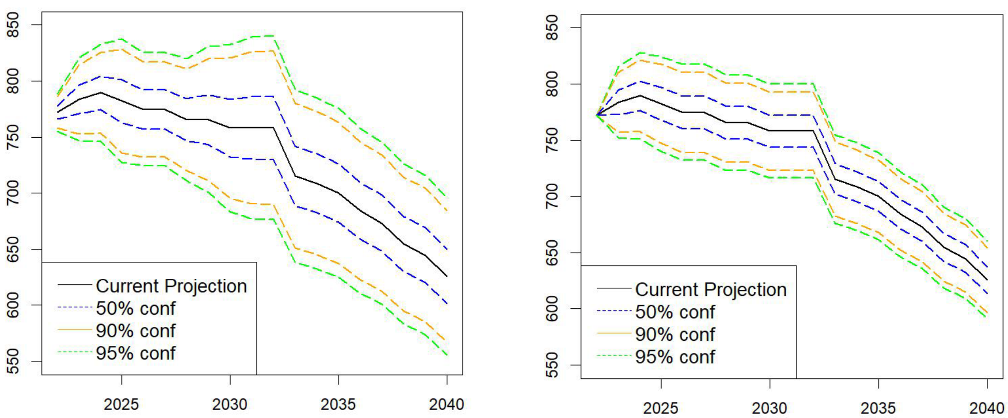
1. Truncate the historical series, using the longest most recent period for which the data appear Gaussian.
2. Apply logarithmic or other transformations to the proportional changes to induce normality before computing standard deviations. (Reverse transform before computing uncertainty bounds.)
3. Compute uncertainty cones using a Gaussian distribution with the standard deviations set to the differences between the Reference case projections and the projections from the farthest AEO side case (Kaack's method SP_1).

As an example of approach (1) above, the available series of annual relative historical changes in nuclear electricity generation (billion kwh) runs from 1957 to 2022 and is non-Gaussian. Figure 5.2.1 shows the G_1 and G_2 uncertainty cones for this series. The G_2 cones are clearly too wide to adjust to the level of the G_1 cones without the use of extreme adjustment factors. If we truncate the series and use only data from 2002 to 2022 (Figure 5.2.1 (b)), the resulting uncertainty cones are much narrower and are the same order of magnitude as the G_1 cones. Although the G_2 cones computed from the truncated series are narrower than the G_1 cones, they can be adjusted using factors that may be generalizable to other series.



(a) (b)

Figure 5.2.1: Uncertainty cones for nuclear net generation (billion kwh), computed by (a) method G_1 and (b) method G_2, using data from 1957 to 2022 for the G_2 cones



(a) (b)

Figure 5.2.2: Uncertainty cones for nuclear net generation (billion kwh), computed by (a) method G_1 and (b) method G_2, using data from 2002 to 2022 for the G_2 cones

Future research will focus on developing the special procedures for non-Gaussian series and on refining the robust adjustment factors given in Table 2 for the G_2 uncertainty cones.

Acknowledgements

The authors gratefully acknowledge the contributions of EIA colleagues Aidan Whitis, Samson Adeshiyan, David Kinyon, and Edward Prokop to this research.

References

Arora, V. (2013), "An evaluation of macroeconomic models for use at EIA," EIA Working Paper available at https://www.eia.gov/workingpapers/.

Kaack, L.H., Apt, J., Morgan, M.G., McSharry, P. (2017). "Empirical prediction intervals improve energy forecasting," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 114, no. 33, pp. 8752-8757, August 15, 2017.

Sherwin, E.D. (2019). "Decisions and Uncertainties in the US Energy System: Electrofuels and Other Applications," Diss. Carnegie Mellon University.

Sherwin, E.D., Henrion, M., and Azevedo, I.M.L. (2018). "Estimation of the year-on-year volatility and the unpredictability of the United States energy system," *Nature Energy*, vol. 3, pp. 341-346, April 2018.

Shlyakhter, A.I.; Kammen, D.M.; Broido, C.L.; and Wilson, R. (1994). "Quantifying the credibility of energy projections from trends in past data: The U.S. energy sector." *Energy Policy Vol. 22*, pp. 119-130.

Williams, W.H. and Goodman, M.L. (1971). "A Simple Method for the Construction of Empirical Confidence Limits for Economic Forecasts," *Journal of the American Statistical Association*, Vol. 66, No. 336, pp.752-754, December 1971.

Gneiting, T. and Raftery, A.E. (2007). "Strictly Proper Scoring Rules, Prediction, and Estimation," *Journal of the American Statistical Association*, Vol. 102, No. 477, pp. 359-379, March 2007.

Appendix A: Supplementary Results

Here we present additional graphics and analytical results. Figure A.1 shows box plots of the p-values of the Shapiro-Wilk normality test for the series in the 2022 AEO Retrospective analysis, by horizon. The null hypothesis of the Shapiro-Wilk test is normality, so we accept as normal any series/horizon combination whose p-value exceeds 0.05 (the red horizontal line in Figure A.1. Based on these results, we conclude that the relative retrospective errors are largely normal and that the use of method G_1 is justified.

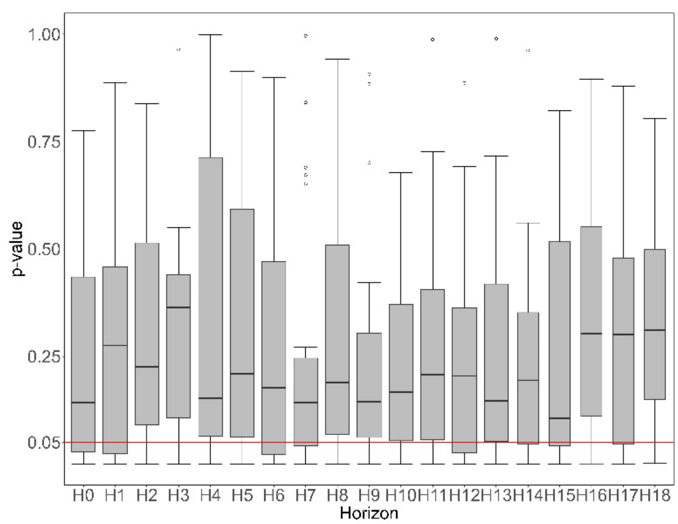


Figure A.1: Box plots of the p-values of the Shapiro-Wilk normality test for the relative projection errors published in the AEO 2022 Retrospective analysis. Based on the test results, we conclude that the use of method G_1 is justified.

For the G_2 uncertainty cones, we evaluate the normality of the relative changes in the historical series. For almost all series, the changes at the lower horizons pass the normality test, as shown in Figures A.2 through A.4. For the higher horizons, the results vary widely. For the series in AEO Table 1, "Total Energy Supply, Disposition, and Prices," normality holds only up to about horizon 15 (Figure A.2). By contrast, for series in AEO Table 4, "Residential Key Indicators and Consumption," normality holds past horizon 50. (Many of these series go back several decades.) For series in AEO Table 8, "Electricity Supply, Disposition, Prices, and Emissions," normality breaks down after about horizon 10. This table includes the solar, wind, and nuclear electricity generation series, for which we're developing special procedures (see Section 5).

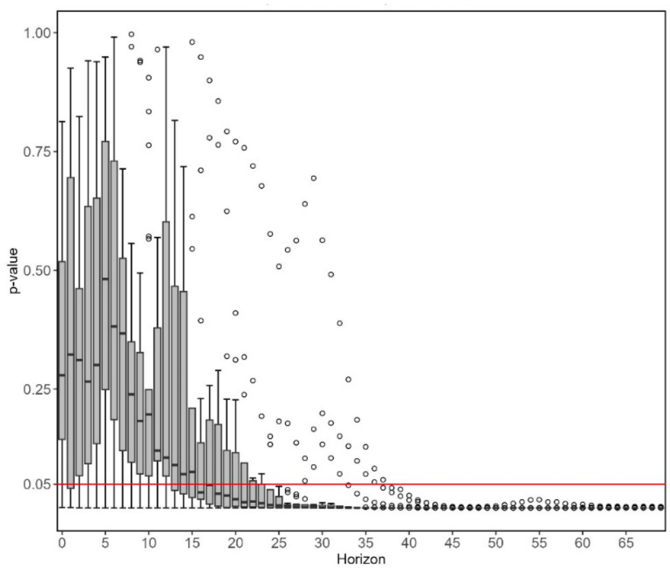


Figure A.2: Box plots of the p-values of the Shapiro-Wilk normality test for the historical relative changes for series in AEO Table 1, "Total Energy Supply, Disposition, and Prices"

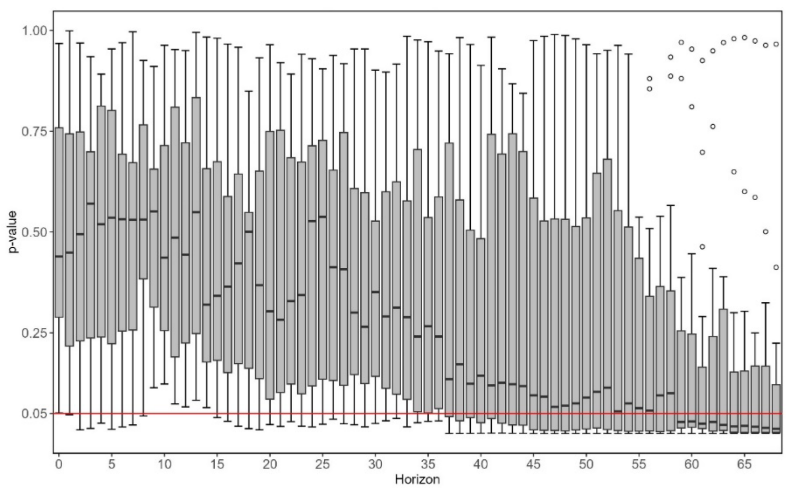


Figure A.3: Box plots of the p-values of the Shapiro-Wilk normality test for the historical relative changes for series in AEO Table 4, "Residential Key Indicators and Consumption"

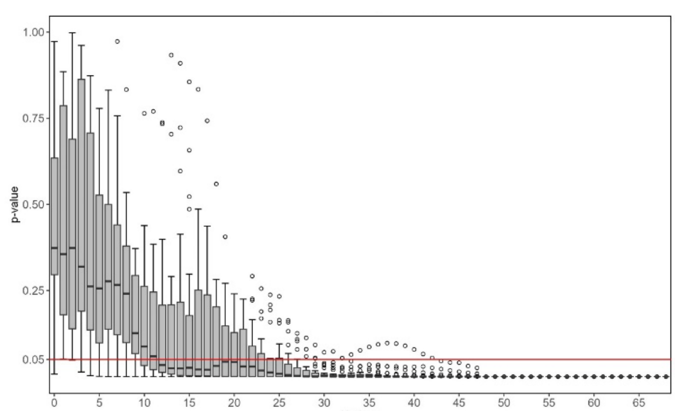


Figure A.4: Box plots of the p-values of the Shapiro-Wilk normality test for the historical relative changes for series in AEO Table 8, "Electricity Supply, Disposition, Prices, and Emissions"

Table A1: Regression Results for Model 1 (EP = Electric Power sector)

Series Description	$\hat{\beta}_0$	$\hat{\beta}_1$	*p*-value, $\hat{\beta}_0$	*p*-value, $\hat{\beta}_1$	Adj. R^2
Prices					
Natural gas price, EP (const. $)	0.07762	0.49731	0.21100	1.646E-06	0.80264
Natural gas price, EP (nom. $)	0.06049	0.58000	0.33364	2.498E-06	0.79070
Coal price, EP (constant $)	-0.08361	0.85323	0.00250	3.806E-10	0.93978
Coal price, EP (nominal $)	-0.07709	1.05971	0.00352	2.804E-10	0.94235
Ave. electricity price (const. $)	0.05255	0.14579	0.01673	0.00346	0.43025
Ave. electricity price (nom. $)	0.03532	0.32532	0.03572	0.00002	0.72230
Imported cost of oil (const. $)	-0.17515	0.89981	0.03458	4.171E-08	0.88263
Imported cost of oil (nom. $)	-0.17795	0.98535	0.05184	2.036E-07	0.85308
Production					
Crude oil production	0.03760	0.44680	0.00741	1.073E-08	0.90319
Dry natural gas production	0.03342	0.36924	0.06209	3.004E-06	0.78519
Coal prod. excl. waste coal	-0.31924	4.27892	4.981E-08	1.358E-11	0.96254
Consumption					
Liquids consumption	0.02400	0.28563	0.05104	0.00007	0.66937
Natural gas consumption	0.04925	0.20930	0.00752	0.00052	0.55951
Coal consumption	-0.15947	2.20070	1.764E-06	4.270E-13	0.97713
Total energy consumption	0.01453	0.29705	0.00696	4.029E-08	0.88320
Residential consumption.	0.02893	0.30245	0.00002	5.795E-08	0.87703
Commercial consumption	0.01430	0.41989	0.00021	2.393E-13	0.97894
Industrial consumption	0.01891	0.37600	0.10358	0.00010	0.64968
Transportation consumption	0.01529	0.45617	0.00451	1.957E-10	0.94522
EP sales excl. direct use	0.00694	0.16405	0.00413	6.233E-15	0.98749
Imports					
Petroleum net imports	0.67264	7.75921	0.86763	0.03651	0.22463
Natural gas net imports	-3.59228	0.53413	0.00150	1.089E-08	0.90299
Electricity generation					
Solar net generation	0.15312	-0.00203	2.491E-10	1.977E-06	0.82155
Wind net generation	0.17137	-0.00003	0.00001	0.16871	0.07433
Hydroelectric net generation	0.07963	-0.04166	0.00001	0.30181	0.01103
Coal net generation	0.02623	0.42862	0.27223	0.00000	0.84917
Natural gas net generation	0.07025	0.03467	0.00005	0.10581	0.12614
Nuclear net generation	0.03660	-0.00003	0.00000	0.06521	0.17914
Macro					
Real GDP (cum. growth)	0.00911	0.00011	0.00007	0.69246	-0.05909
Energy-related CO2 emissions	-0.02329	1.20257	1.540E-07	2.171E-17	0.99442
Energy intensity	0.01250	1.01845	0.00003	1.025E-14	0.98657

Table A2: Regression Results for Model 2 (EP = Electric Power sector)

	$\hat{\gamma}$	***p*-value,** $\hat{\gamma}$	**Adj.** R^2
Prices			
Natural gas price, EP (const. $)	0.57527	5.372E-14	0.97758
Natural gas price, EP (nom. $)	0.65253	4.072E-14	0.97839
Coal price, EP (constant $)	0.65718	5.758E-15	0.98335
Coal price, EP (nominal $)	0.83320	2.687E-15	0.98496
Ave. electricity price (const. $)	0.24919	2.225E-09	0.90785
Ave. electricity price (nom. $)	0.43601	3.678E-12	0.96065
Imported cost of oil (const. $)	0.70942	2.427E-14	0.97983
Imported cost of oil (nom. $)	0.76966	7.719E-14	0.97647
Production			
Crude oil production	0.55519	8.138E-15	0.98256
Dry natural gas production	0.45807	1.194E-11	0.95398
Coal prod. excl. waste coal	2.05400	5.651E-10	0.92317
Consumption			
Liquids consumption	0.38374	1.080E-10	0.93833
Natural gas consumption	0.33781	3.640E-09	0.90164
Coal consumption	1.56769	7.508E-13	0.96815
Total energy consumption	0.37534	2.659E-13	0.97226
Residential consumption.	0.47000	2.322E-12	0.96299
Commercial consumption	0.49083	2.675E-17	0.99186
Industrial consumption	0.48872	2.859E-11	0.94831
Transportation consumption	0.54169	2.076E-15	0.98546
EP sales excl. direct use	0.17797	1.100E-18	0.99468
Imports			
Petroleum net imports	8.27558	0.00002	0.68892
Natural gas net imports	0.37840	2.261E-09	0.90766
Electricity generation			
Solar net generation	0.00155	0.03452	0.23008
Wind net generation	0.00009	0.00214	0.46592
Hydroelectric net generation	0.22140	0.00001	0.76236
Coal net generation	0.47760	3.687E-12	0.96890
Natural gas net generation	0.13696	2.651E-06	0.78894
Nuclear net generation	0.00005	0.14997	0.08089
Macro			
Real GDP (cum. growth)	0.00162	2.196E-11	0.95009
Energy-related CO2 emissions	0.99339	1.463E-17	0.99249
Energy intensity	1.19476	1.189E-19	0.99605

Regression Diagnostics for the Effect of Category on $\hat{y}$

```
Call:
lm(formula = Gamma ~ -1 + Category, data = GammaTest)

Residuals:
    Min      1Q  Median      3Q     Max
-3.9486 -0.1680 -0.0463  0.1292  3.9486

Coefficients:
                     Estimate Std. Error t value Pr(>|t|)
Categoryconsumption    0.5371     0.3962   1.356    0.187
Categorygeneration     0.1396     0.4852   0.288    0.776
Categoryimports        4.3270     0.8405   5.148 2.54e-05
Categorymacro          0.7299     0.6862   1.064    0.298
Categoryprice          0.6103     0.4202   1.452    0.159
Categoryproduction     1.0224     0.6862   1.490    0.149


Multiple R-squared:  0.5755,     Adjusted R-squared:  0.4736
F-statistic: 5.648 on 6 and 25 DF, p-value: 0.0008072
```

Appendix B: Technical Details

Formulas for Uncertainty Metrics

As in Kaack et al. (2017), for series other than prices, we define $\varepsilon_{rel} = \frac{\hat{y}-y}{y}$, where $\hat{y}$ is the AEO Reference case projection of the historical value y, and use the approximation

$$s_r \equiv sd(\varepsilon_{rel}) = sd\left(\frac{\hat{y}-y}{y}\right) \quad \text{(B1)}$$

to compute the G_1 uncertainty cones for non-price series. We use the approximation $s_r = sd\left(\frac{\hat{y}-y}{y}\right)$ for the ratio $\frac{sd(\hat{y}-y)}{y}$ because, for each horizon, we pool the relative errors across time periods, so the historical values y (in the denominator) vary across the ratios. When we compute the error bounds, we multiply s_r by the current projection $\hat{y}$ (see equation B3).

For price series, which we analyze in log scale, we define $\varepsilon_{log} = \ln\hat{y} - \ln y$ and, for each horizon, consider the standard deviation

$$s_l \equiv sd\left(\varepsilon_{log}\right) = sd(\ln\hat{y} - \ln y). \quad \text{(B2)}$$

With approximation (1), we compute the upper and lower G_1 Gaussian confidence bounds $b_{U,Q}$ and $b_{L,Q}$ for confidence level Q for a projected series $\hat{y}$ other than prices as

$$b_{U,Q} \approx \hat{y} + \left(a_{U,Q} \times \hat{y} \times s_r\right), \quad \text{(B3)}$$

and

$$b_{L,Q} \approx \hat{y} + \left(a_{L,Q} \times \hat{y} \times s_r\right), \quad \text{(B4)}$$

where $a_{U,P}$ and $a_{L,P}$ represent the appropriate Gaussian quantiles. Note that, in this context, $\hat{y}$ is a current projection for a future year. For price series, we compute the Gaussian confidence bounds as

$$b_{U,Q,P} \approx \exp\left[\ln\hat{y} + \left(a_{U,Q} \times s_l\right)\right] = \hat{y} \times \exp\left(a_{U,Q} \times s_l\right), \quad \text{(B5)}$$

and

$$b_{L,Q,P} \approx \exp[\ln \hat{y} + (a_{L,Q} \times s_l)] = \hat{y} \times \exp(a_{L,Q} \times s_l). \qquad \text{(B6)}$$

The relevant quantiles $a_{U,Q}$ and $a_{L,Q}$ from the Gaussian density are as follows:

Q	$a_{U,Q}$	$a_{L,Q}$
0.50	0.674	-0.674
0.90	1.645	-1.645
0.95	1.96	-1.96

To compute the G_2 uncertainty cones, we follow the same procedure as for the G_1 cones, except that we substitute relative historical changes $\varepsilon_{rel,t} = \frac{y_{t_2}-y_{t_1}}{y_{t_1}}$ for the relative projection errors $\varepsilon_{rel} = \frac{\hat{y}-y}{y}$ for non-price series. Similarly, for price series, we substitute log-scaled historical changes $\varepsilon_{log,t} = \ln y_{t_2} - \ln y_{t_1}$ for the log-scaled relative projection errors $\varepsilon_{log} = \ln \hat{y} - \ln y$.

To form the centered empirical error bounds (method NP_2), we first compute the centered relative (or log) errors for each horizon:

$$\varepsilon_{rel,ctr} \equiv \varepsilon_{rel} - median(\varepsilon_{rel}), \qquad \text{(B7)}$$

and

$$\varepsilon_{log,ctr} \equiv \varepsilon_{log} - median(\varepsilon_{log}), \qquad \text{(B8)}$$

where the medians are taken over all ε_{rel} or ε_{log} associated with a horizon. We then compute the quantiles of the centered error distributions. Because this is a two-sided distribution, we compute the following quantiles for the upper and lower bounds:

Q	$c_{U,Q}$	$c_{L,Q}$
0.50	0.750	0.250
0.90	0.950	0.050
0.95	0.975	0.025

Given the centered empirical quantiles, we compute the centered empirical error bounds:

$$D_{U,Q} \approx \hat{y} + (C_Q \times \hat{y}), \qquad \text{(B9)}$$

and

$$D_{L,Q} \approx \hat{y} - (C_Q \times \hat{y}), \qquad \text{(B10)}$$

where C_Q represents the appropriate empirical quantile of the $\varepsilon_{rel,ctr}$. For price series, we compute the empirical confidence bounds as

$$D_{U,Q,P} \approx \exp[\ln \hat{y} + C_{Q,P}] = \hat{y} \times e^{C_{Q,P}}, \qquad \text{(B11)}$$

and

$$D_{L,Q,P} \approx \exp[\ln \hat{y} - C_{Q,P}] = \hat{y} \times e^{-C_{Q,P}}, \qquad \text{(B12)}$$

where $C_{Q,P}$ represents the appropriate quantile of the centered empirical distribution.

Continuous Ranked Probability Score (CRPS)

For each forecast horizon H, and each forecast, let F_t be the cumulative distribution function (CDF) of the forecast density, let n_H be the number of pairs of observations and forecasts H years apart (i.e., having forecast horizon H), let ξ_t be the forecast error for year t, and let ϵ_t be a point of the predictive error distribution. The CRPS is defined as

$$CRPS_H(F_t, \epsilon) = \frac{1}{n_H}\sum_{t=1}^{n_H}\int_{-\infty}^{\infty}[F_t(\epsilon_t) - \mathbf{1}(\epsilon_t \geq \xi_t)]^2 d\epsilon_t, \quad \text{(B13)}$$

where $\mathbf{1}(\epsilon_t \geq \xi_t)$ is a step function that is 0 for all real values less than ξ_t and 1 for all real values greater than or equal to ξ_t. The CRPS is a scoring function that essentially compares the CDF of the density forecast with a step function that represents the CDF of the difference between an AEO projection and its true value (based on a retrospective analysis). It averages the squared differences between these CDFs across each forecast probability distribution (the integral) and across all forecast-observation pairs with horizon H (the sum). A lower CRPS indicates that the forecast densities are closer to the step functions, on average, and therefore closer to representing the true error distribution. Because the step function is infinitely sharp, the CRPS is a measure of both calibration and sharpness (favoring narrower densities).