Empirical Prediction Intervals Improve Energy Forecasting

Lynn H. Kaack, Jay Apt, M. Granger Morgan and Patrick McSharry

Department of Engineering and Public Policy, Carnegie Mellon University, 5000 Forbes Ave, Pittsburgh, PA 15213

Overview

While many forecasters are moving towards generating probabilistic predictions, energy forecasts typically still consist of point projections and scenarios without associated probabilities. Empirical density forecasting methods provide a probabilistic amendment to existing point forecasts. Here we lay the groundwork for evaluating the performance of these methods in the data-scarce setting of long-term forecasts. Results can give policy analysts and other users confidence in estimating forecast uncertainties with empirical methods.

Density forecast

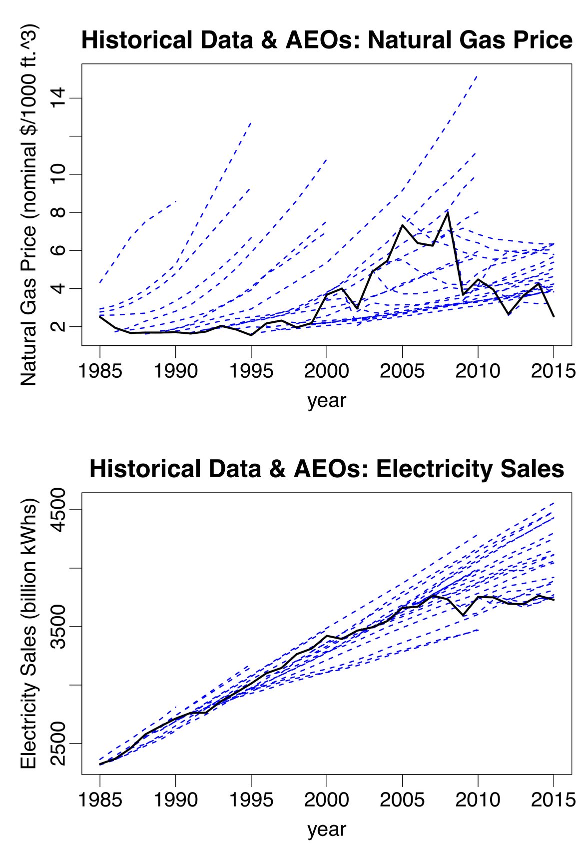

Figure 1: The historical values and AEO projections for natural gas wellhead prices. Black indicates the historical averages and blue the AEOs 1982-2016.

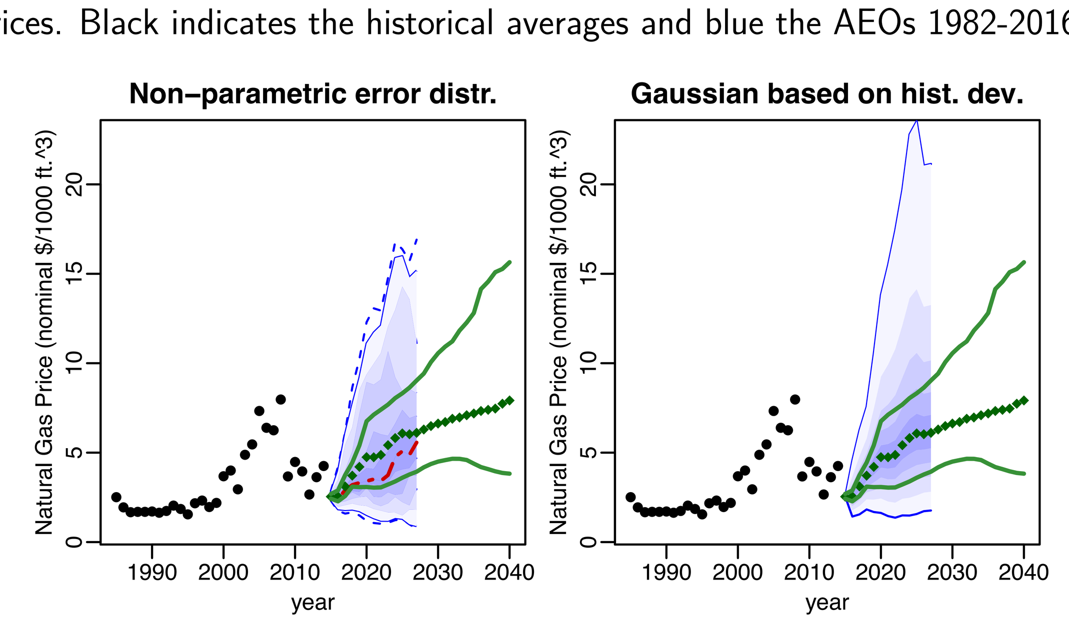

Figure 2: Density forecasts for natural gas prices in nominal $. (A) NP_1 (B) G_2, which tested to be the better estimate. Historical values in black, the AEO 2016 reference case and scenarios in green and the density forecast in blue shaded areas, corresponding to the percentiles 2, 10, 20, 30, ..., 80, 90, 98.

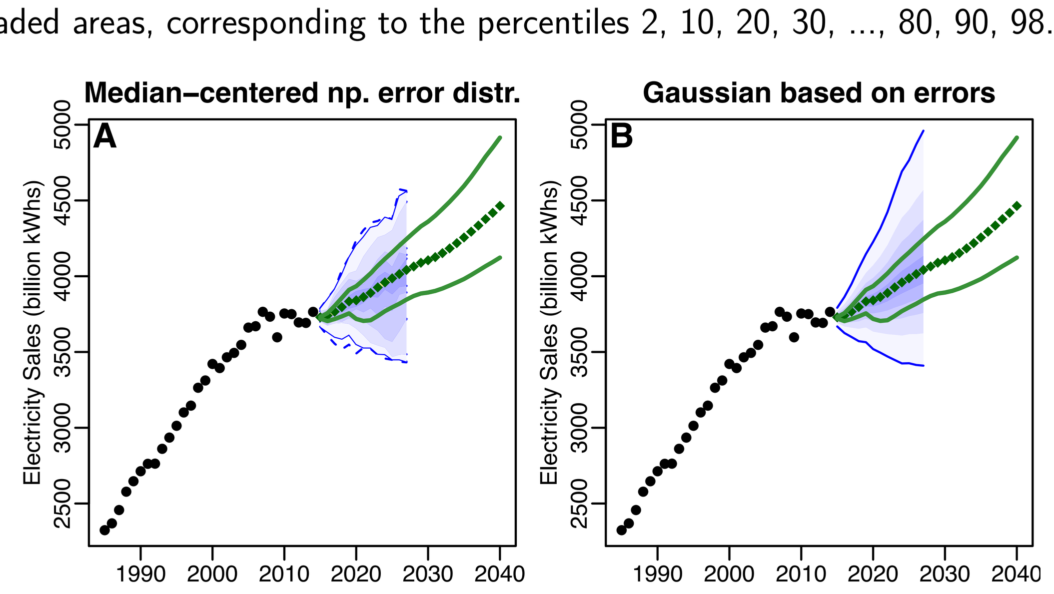

Figure 3: (A) NP_2, the median or bias now coincides with the AEO reference case. (B) G_1 was the best forecast over the test range.

We compare a number of methods

Point forecasting methods:

- *AEO reference case*: We treat the AEO reference case as a point forecast.
- *Median errors* (NP_1): The median of the EPI with a non-parametric distribution of the errors (NP_1), computed as the reference case adjusted by the median of past forecasting errors.
- *Persistence*: Persistence refers to a constant forecast equal to the last observation. Here, we use the forecasted value at $H = 0$ as the last observation.
- *Simple linear model:* A simple linear regression with time as the predictor and an optimal moving window of the last 7 historical observations.

Density forecasting methods

- NP_1: EPI with a non-parametric distribution of the forecasting errors and a median different to the reference case. This method was originally published by [1].
- NP_2: EPI with a non-parametric error distribution, which is centered such that the AEO reference case is the best estimate forecast.
- G_1: A Gaussian distribution with the SD of the past errors and mean and median of $\epsilon = 0$.
- G_2: Gaussian with a SD based on a sample of all relative deviations between two historical data points which are H steps apart. Mean and median are $\epsilon = 0$.
- S: Ensemble forecast of the reference case and the highest and lowest scenario projection.
- SP: A Gaussian distribution with the distance to the farthest scenario as 1 SD (SP_1) and a uniform distribution between the envelope scenarios (SP_2).

Table 1: Empirical density forecasting methods compared

Method	Parametric	Based on	Median ctrd.
NP1	no	forecast errors	no
NP2	no	forecast errors	yes
G1	yes	forecast errors	yes
G2	yes	hist. deviations	yes

Important Result

We report probabilistic uncertainties for 18 core quantities of the AEO 2016 projections. Our work frames how to produce, evaluate and rank probabilistic forecasts in this setting.

Mathematical section: error transformation and validation scores

Relative error: $\epsilon_{rel} = \frac{\hat{y}-y}{y} = \frac{\hat{y}}{y} - 1$

For price quantities: drawing an analogy to logarithmic returns we modify to yield the **logarithmic error** $\epsilon_{log} = \ln(1+\epsilon_{rel}) = \ln\left(\frac{\hat{y}}{y}\right) = \ln\hat{y} - \ln y$.

For point forecast performance:

$$MAPE_H = \frac{1}{n_H}\sum_{t=1}^{n_H} |\xi_{rel,H,t}| = \frac{1}{n_H}\sum_{t=1}^{n_H} \left|\frac{\hat{y}_{H,t} - y_{H,t}}{y_{H,t}}\right|,$$

and $MALE_H = \frac{1}{n_H}\sum_{t=1}^{n_H} |\ln\hat{y}_{H,t} - \ln y_{H,t}|$, with horizon H, $\hat{y}$: forecast, y: actual observation.

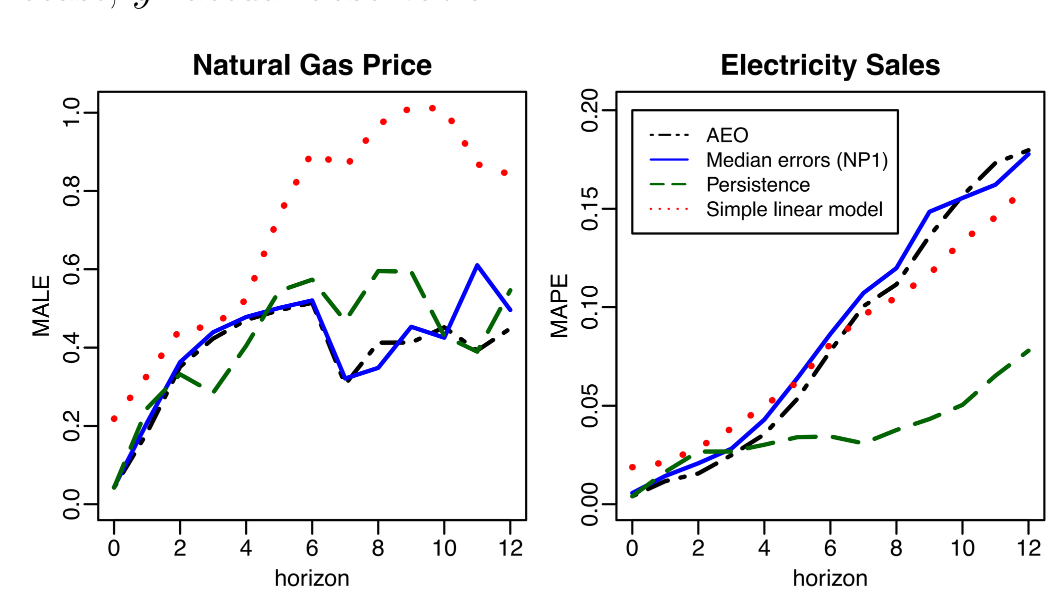

Figure 4: The mean absolute percentage or log-error (MAPE/MALE) for the test range 2003-2014. We see that for natural gas prices (in nominal $), the median of NP_1 performs similarly to the AEO reference case. For electricity sales, the reference case outperforms the median for nearly every horizon.

The continuous ranked probability score (CPRS) for density forecasts:

$$CRPS_H(F,\epsilon) = \frac{1}{n_H}\sum_{t=1}^{n_H}\int_{-\infty}^{\infty} \left(F_t(\epsilon_t) - I(\epsilon_t \geq \xi_t)\right)^2 d\epsilon_t$$

similar to [2]. ϵ_t is a point of the predictive error distribution, while ξ_t is the forecast error of the observation. The CRPS compares the CDF of the density forecast with the CDF of an observation, a step function $I(\epsilon_t \geq \xi_t)$.

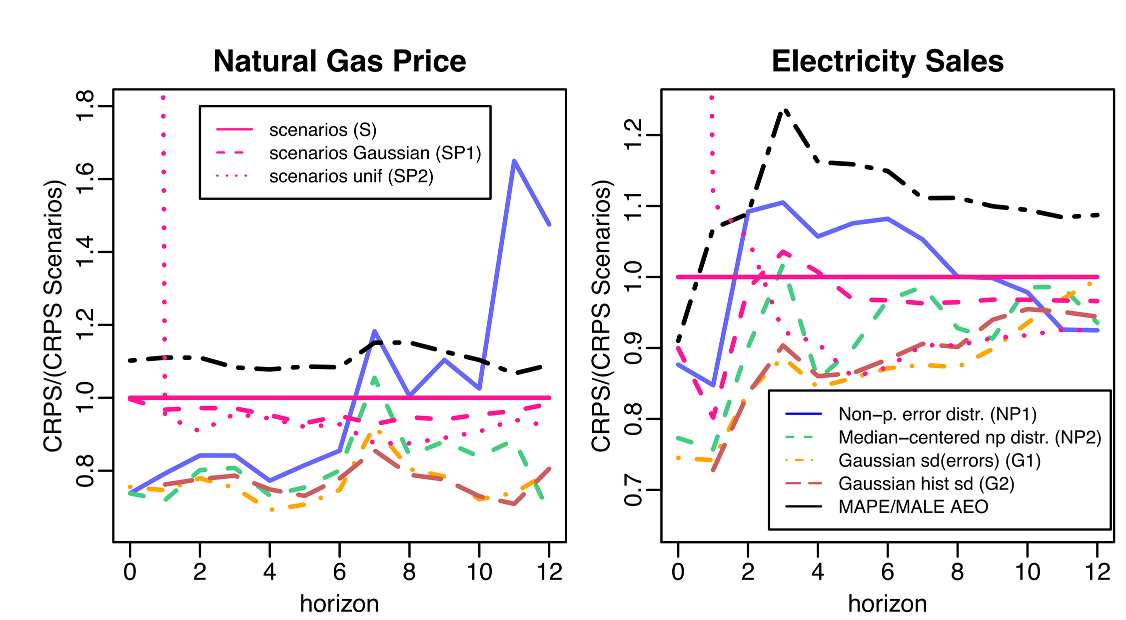

Figure 5: Relative improvement of the methods with respect to the envelope scenarios for the test range 2003-2014. Values are plotted as fraction of the CRPS of the scenario ensemble (S). A normalized CRPS lower than 1.0 corresponds to a better density forecast. SP_1 corresponds to a normal distribution with the scenario range as 1 SD, and SP_2 is a uniform PDF between the envelope scenarios.

Results

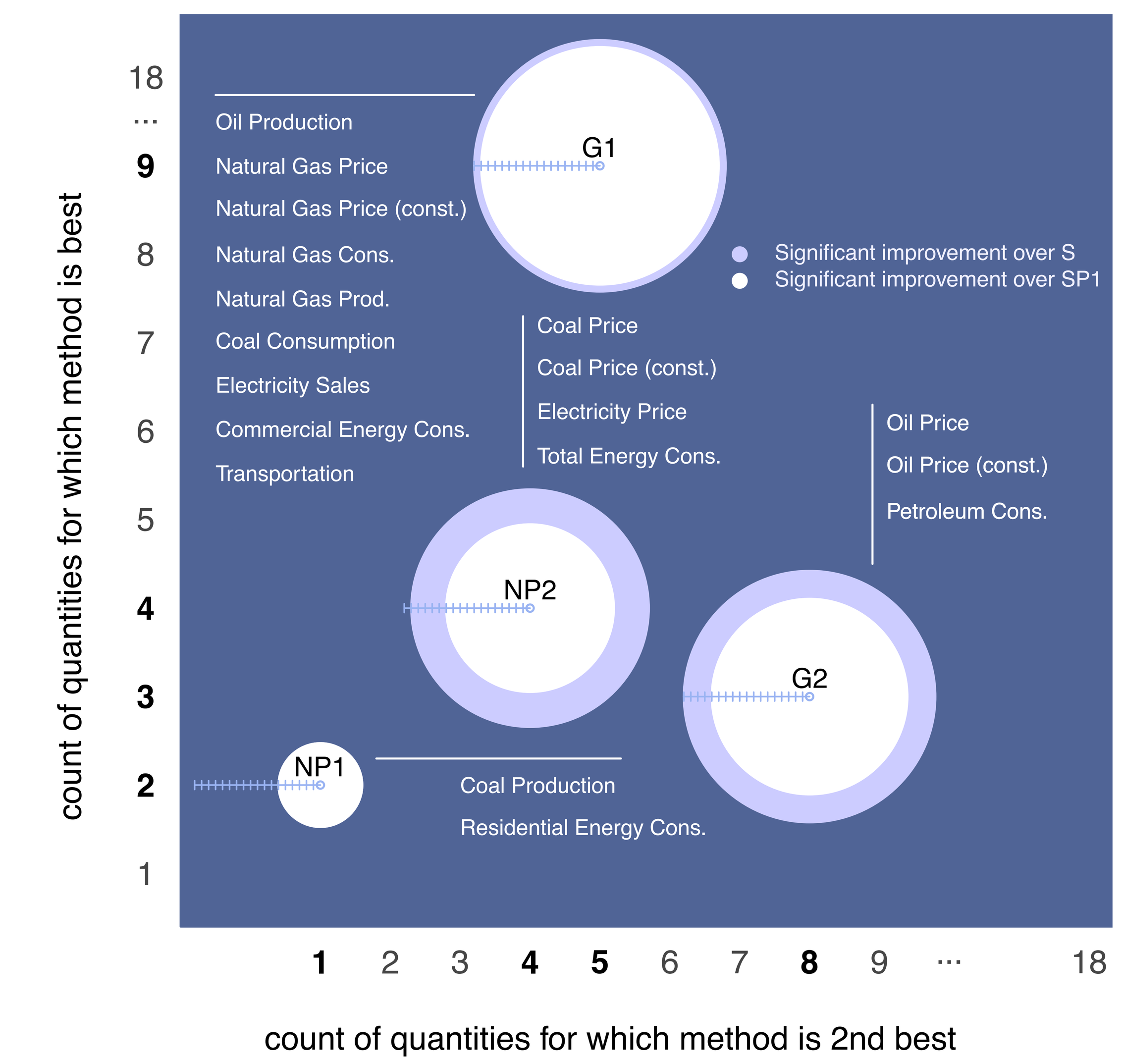

Figure 6: Graphical summary of the two evaluation criteria: best method for number of quantities out of 18 (axes), and significant improvement over scenarios (circle radius).

Conclusion

- There are empirical methods available for estimating the uncertainty around the AEO reference case, which have proven to be significantly more accurate over the past decade than the scenarios of the AEO.
- Gaussian distribution based on past errors (G_1) has good performance over different quant..
- We recommend that the EIA include the standard deviation of forecast errors in their retrospective reports
- Representing the exact error distribution does not need to provide the better out-of-sample forecast.
- EIA's forecast bias is in most cases not consistent and a bias-corrected reference case is not advised.

References

[1] W H Williams and M L Goodman.
A simple method for the construction of empirical confidence limits for economic forecasts.
Journal of the American Statistical Association, 66(336):752–754, 1971.

[2] Tilmann Gneiting and Adrian E Raftery.
Strictly proper scoring rules, prediction, and estimation.
Journal of the American Statistical Association, 102(477):359–378, 2007.

Acknowledgements

We thank Evan D. Sherwin, Inês L. Azevedo, Cosma R. Shalizi, Alexander L. Davis, Stephen E. Fienberg, and Max Henrion for their advice and assistance. We thank the EIA for hosting a presentation and discussion about this work. Funding by Electric Power Research Institute (EPRI) and by CEDM.