A Retrospective Approach to Quantifying Uncertainty in Traffic Forecasts

Jawad Mahmud Hoque, University of Kentucky

Greg Erhardt, University of Kentucky

David Schmitt, Connetics Transportation Group

Mei Chen, University of Kentucky

Marty Wachs, University of California, Los Angeles

Accuracy and Uncertainty

Accuracy

Closeness of observation and measurement or estimate

Retrospective evaluation of forecast quality

Comparison of actual traffic and forecasted traffic

Uncertainty

Estimate of the accuracy. Range in which the real value lies

Prospective modification of forecasts to ensure quality and reliability

Range of values possible for actual traffic

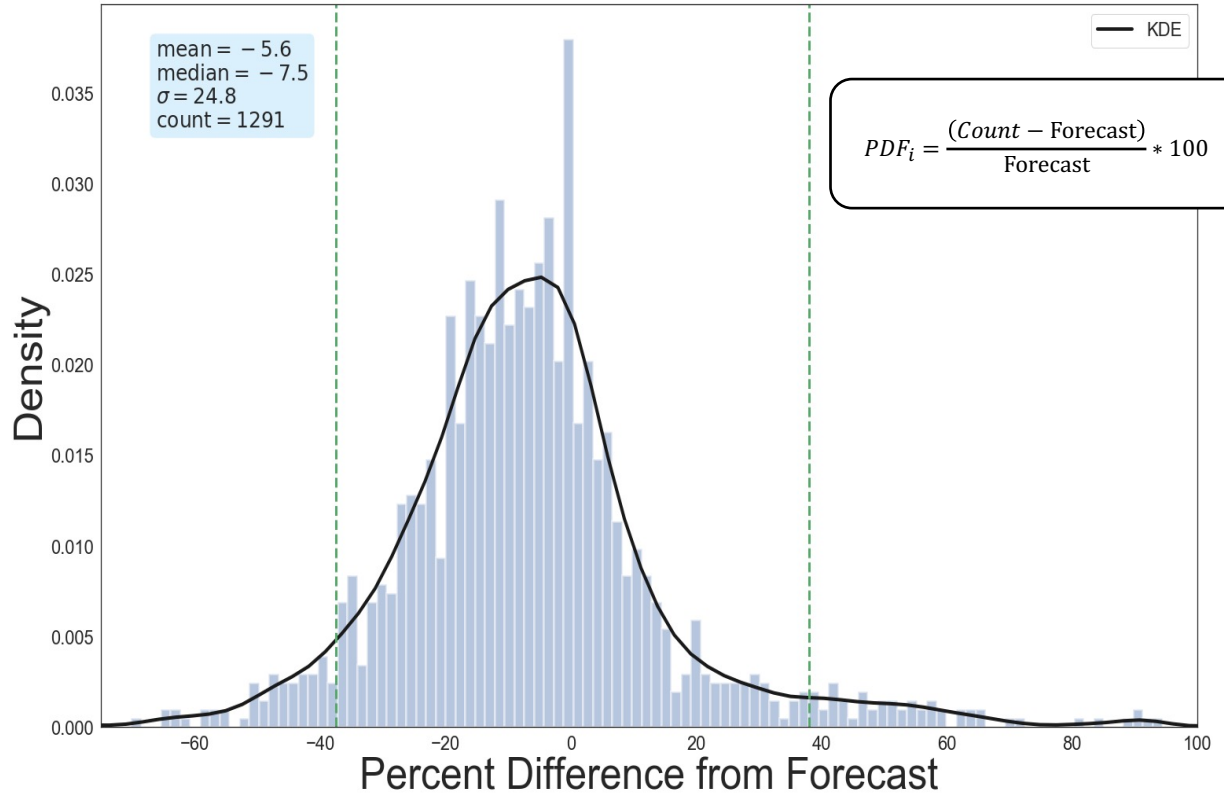
Traffic Forecast Accuracy

National Cooperative Highway Research Program (NCHRP) Project 934 Database on Traffic Forecast Accuracy

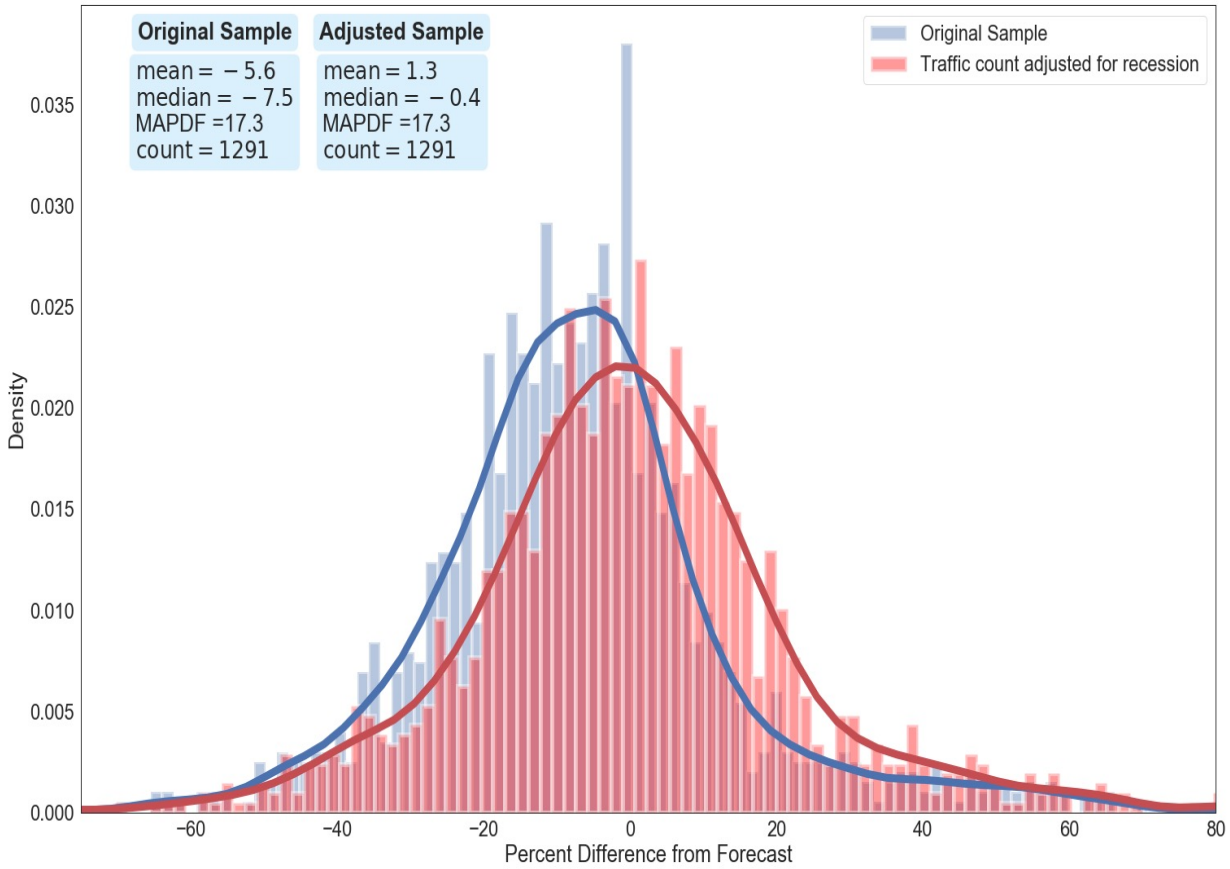
Mean Percent Deviation from Forecast of -5.73%

Mean Absolute Percent Deviation from Forecast of 17.29%

5th percentile is -37% and 95th percentile is +38%. 90% of the forecasts fall in between.

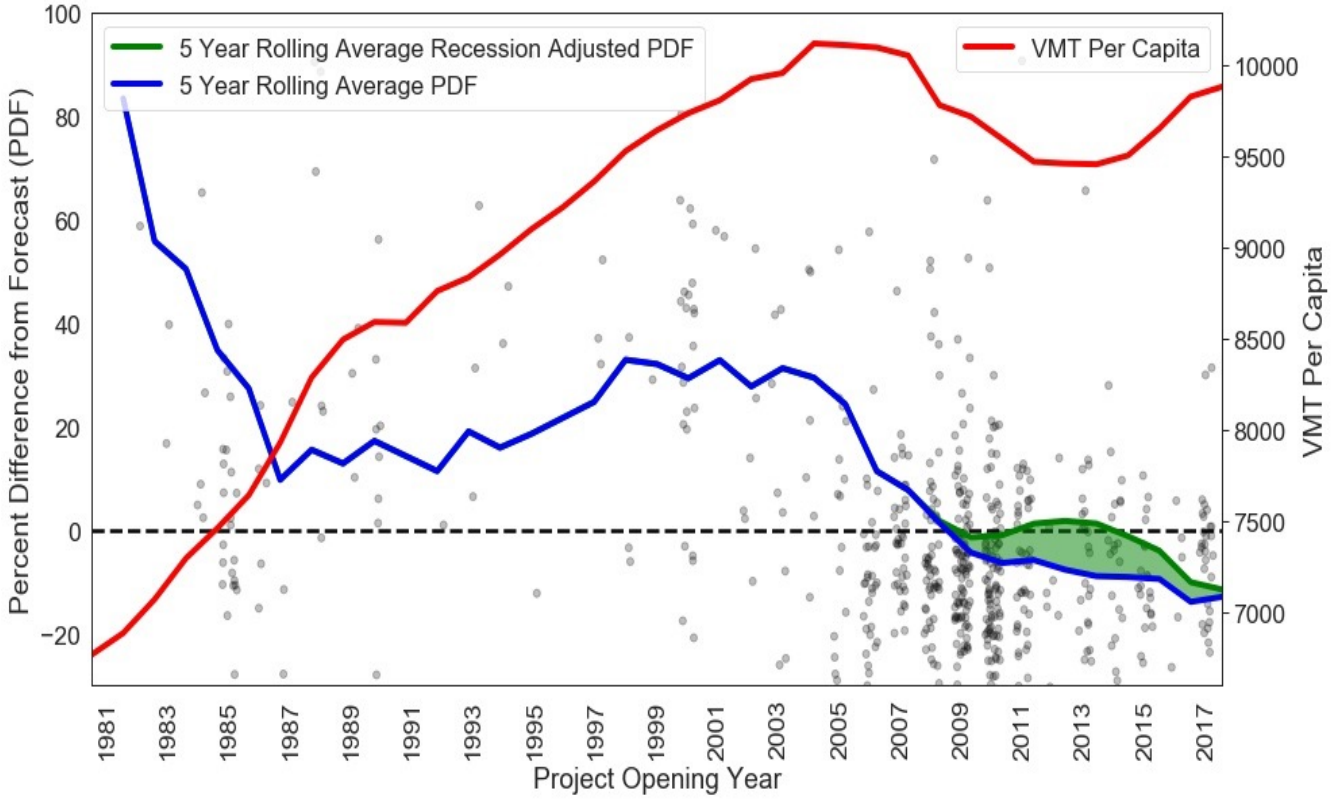


Effect of the Great Recession on Forecast Accuracy



- Traffic would be 1% greater on average, rather than 6% lower, than the forecast if we adjust for higher unemployment during the post-recession years (2008 to 2014).

Effect of the Great Recession on Forecast Accuracy



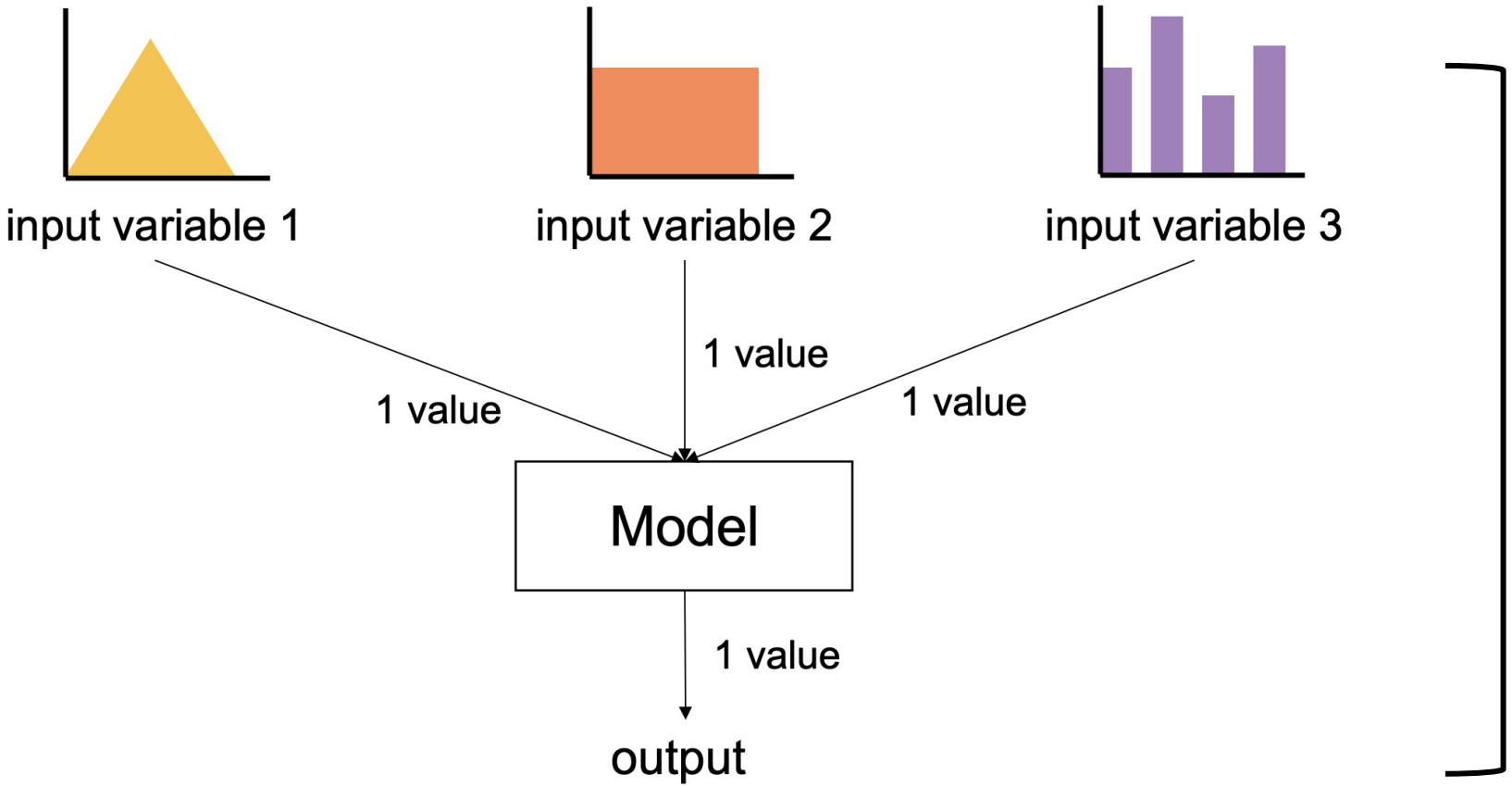
- While VMT per capita was increasing, counted traffic volumes were higher than forecast.
- But after VMT per capita peaks, counted traffic volumes were lower than forecast.
- Economic and fuel price changes determine much of the VMT change. Those same factors may also explain changing traffic forecast accuracy.

Implications

- Forecasts are getting better over the years.
- Higher volume roads, higher functional classes, shorter time spans, and the use of travel models all improved accuracy.
- Forecast accuracy is affected by macro-economic conditions in the project opening year
  - The Great Recession causing a systemic shift in accuracy
- Forecasts may not capture larger VMT trends

Acknowledge Uncertainty By Providing A Range of Expected Traffic

Sensitivity Testing/Scenario Analysis



Repeat with Extremities in input

Limitations

Assumptions about the range of inputs.

Uncertainty in the input data propagates through the model (Zhao and Kockelman, 2002)

Much higher run time on an already time-intensive process.

An Alternate Method

Create uncertainty envelopes around forecasts using empirical evidence of past accuracy

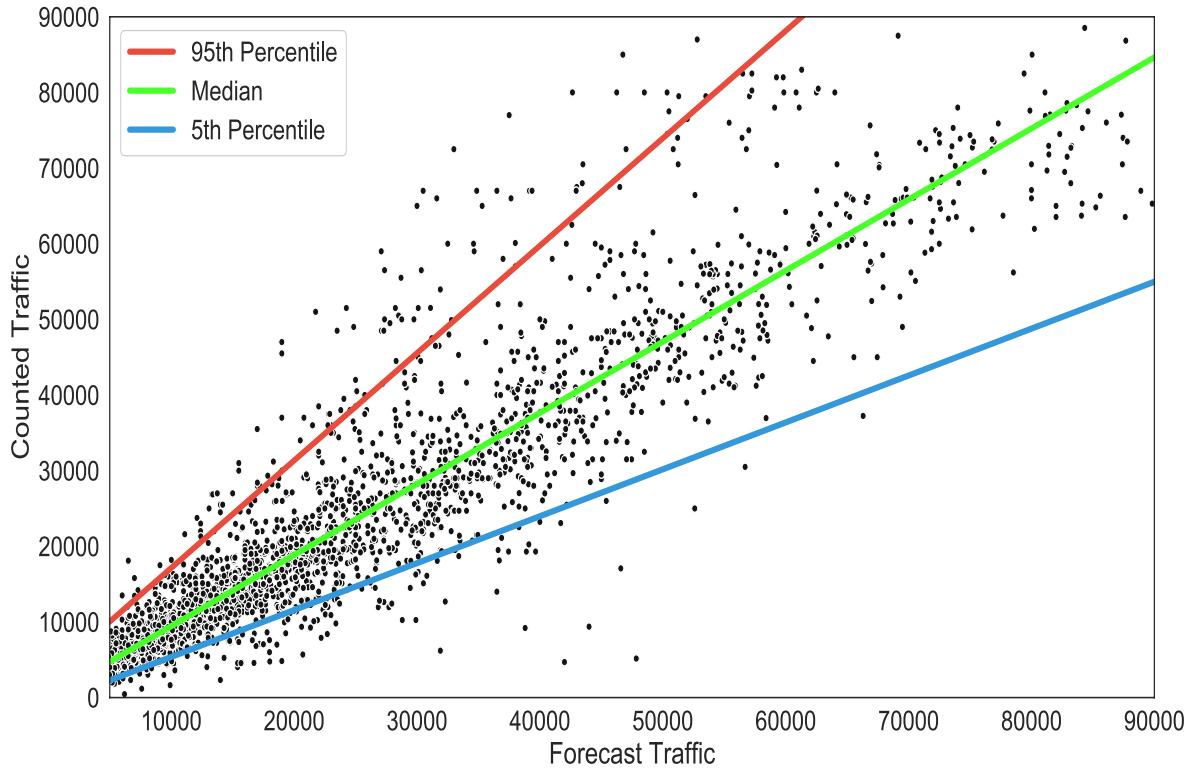
- Inspired by the principle of Reference Class
  - Using the base-rate and distribution results from similar situations in the past to adjust forecasts.
- Will consider the spread of the variables inducing bias

Quantile Regression – A method to both measure accuracy and estimate uncertainty envelopes

Draw line through the middle of the cloud: **regression**.

Draw a line along the edge of the cloud: **quantile regression**.

Quantifying uncertainty is as simple as inputting values in a spreadsheet and drawing lines.



Measuring accuracy and estimating uncertainty windows using Quantile Regression

Model Form

$$y_i = \alpha + \beta \hat{y}_i + \gamma X_i \hat{y}_i + \varepsilon_i$$

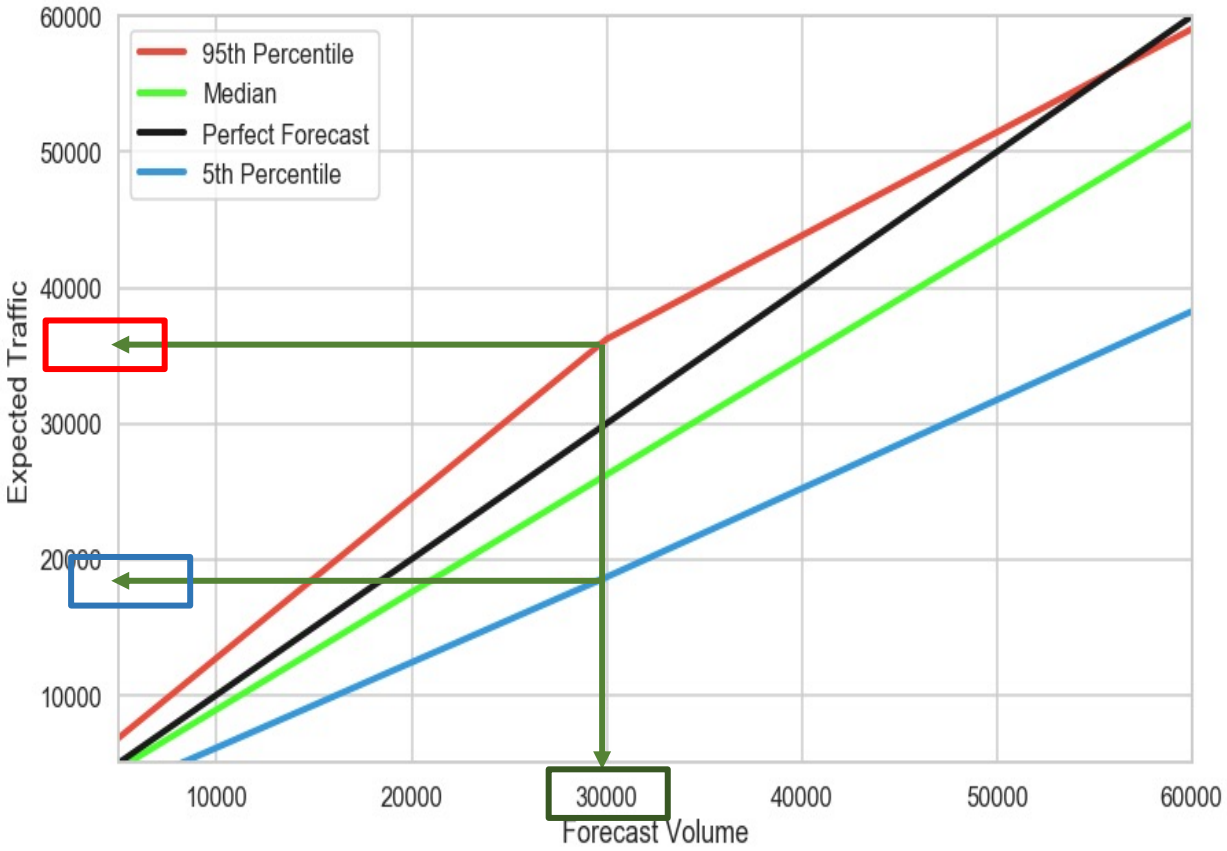
- Multiplicative effect instead of additive
- Estimate separate α, β and γ for different percentile values (95^{th}, 80^{th}, 50^{th}, 20^{th}, 5^{th}).
- Coefficients signify the effect of the explanatory variables on different percentile values of actual observation.
- Example, coefficient of -0.25 on unemployment rate on the 95^{th} percentile model means with each unit increase in unemployment rate, the 95^{th} percentile actual traffic value decreases by 0.25 units.

Model Estimation Results

	5th Percentile	50th Percentile	95th Percentile
Pseudo R-Squared	0.475	0.739	0.83
	Coef.	Coef.	Coef.
Overall Distribution			
Intercept (α)	-182.26	255.55	976.78
Forecast Volume (β)*	0.705	0.891	1.254
Forecast Volume in excess of 30,000 ADT	0.024	-0.004	-0.413
Descriptive Variables			
Time span (years)	0.006	0.008	0.02
Unemployment rate in the year forecast was produced (%)	-0.006	0.002	0.01
Binary Variables			
Functional Class (Reference class = Freeways)			
Major or minor arterials	-0.15	-0.062	-0.116
Collectors and local roads	-0.212	-0.126	-0.321
Project Type (Reference class = Existing Road)			
New road	0.093	-0.008	-0.09
Forecast Method (Reference class = traffic count trend, population growth rate, or professional judgment)			
Travel demand model	0.068	-0.008	-0.101
Year Forecast Produced (Reference class = 2010 or later)			
Years before 2010	-0.007	0.0002	0.003

Uncertainty Envelope-Example 1

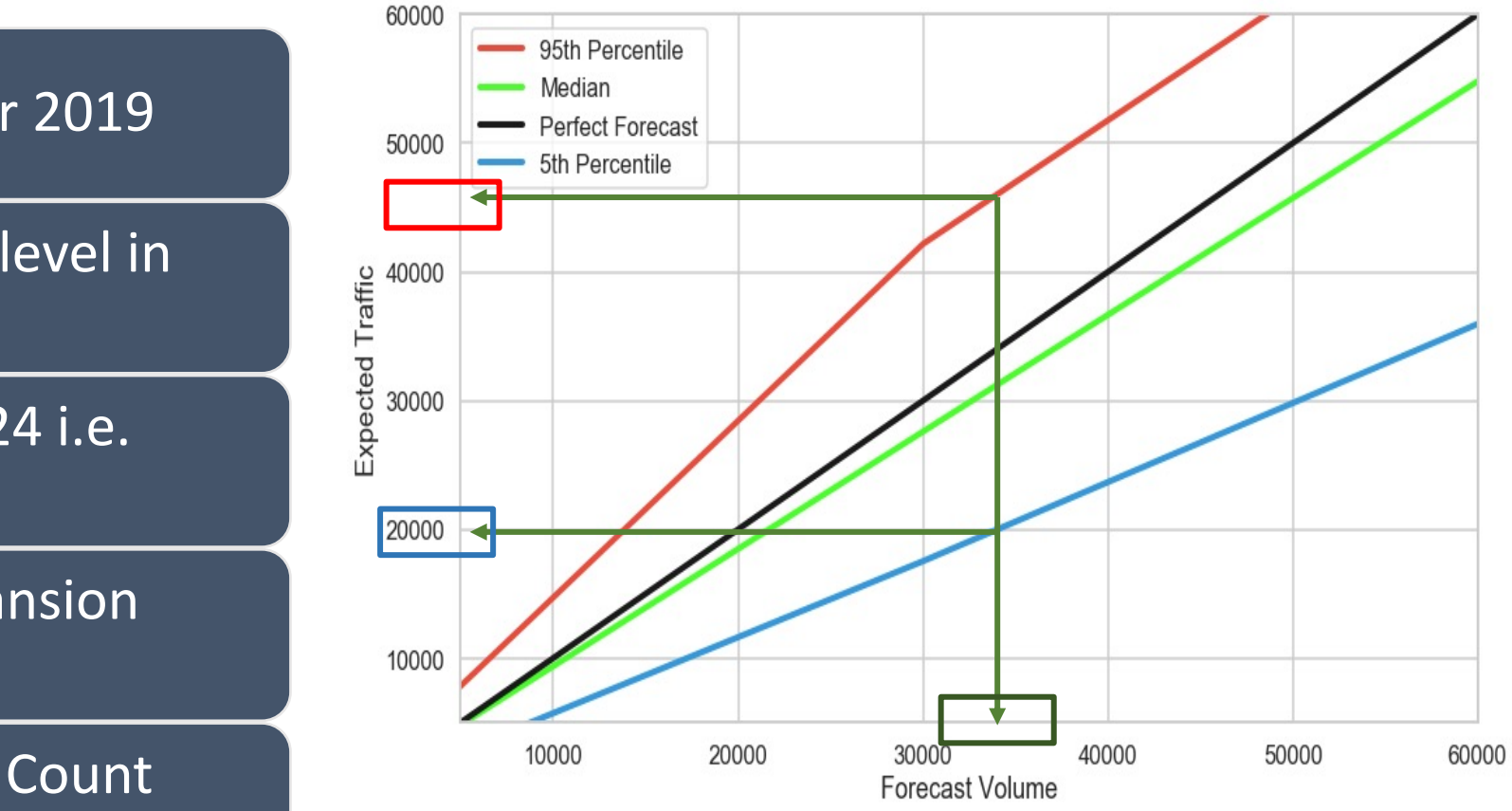
- Forecast produced in the year 2019
- Unemployment rate at State level in 2019 is 4%
- Forecasting the traffic for 2024 i.e. forecast horizon of 5 years
- The project is a capacity expansion project on a Minor Arterial
- Forecast is done using a travel demand model.



For a Forecast of 30,000 the 5th and 95th percentile value of the expected traffic are 19,000 and 36,000 respectively

Uncertainty Envelope-Example 2

- Forecast produced in the year 2019
- Unemployment rate at State level in 2019 is 4%
- Forecasting the traffic for 2024 i.e. forecast horizon of 10 years
- The project is a capacity expansion project on an Arterial
- Forecast is done using Traffic Count Trend



For a Forecast of 35,000 the 5th and 95th percentile value of the expected traffic are 20,000 and 46,000 respectively

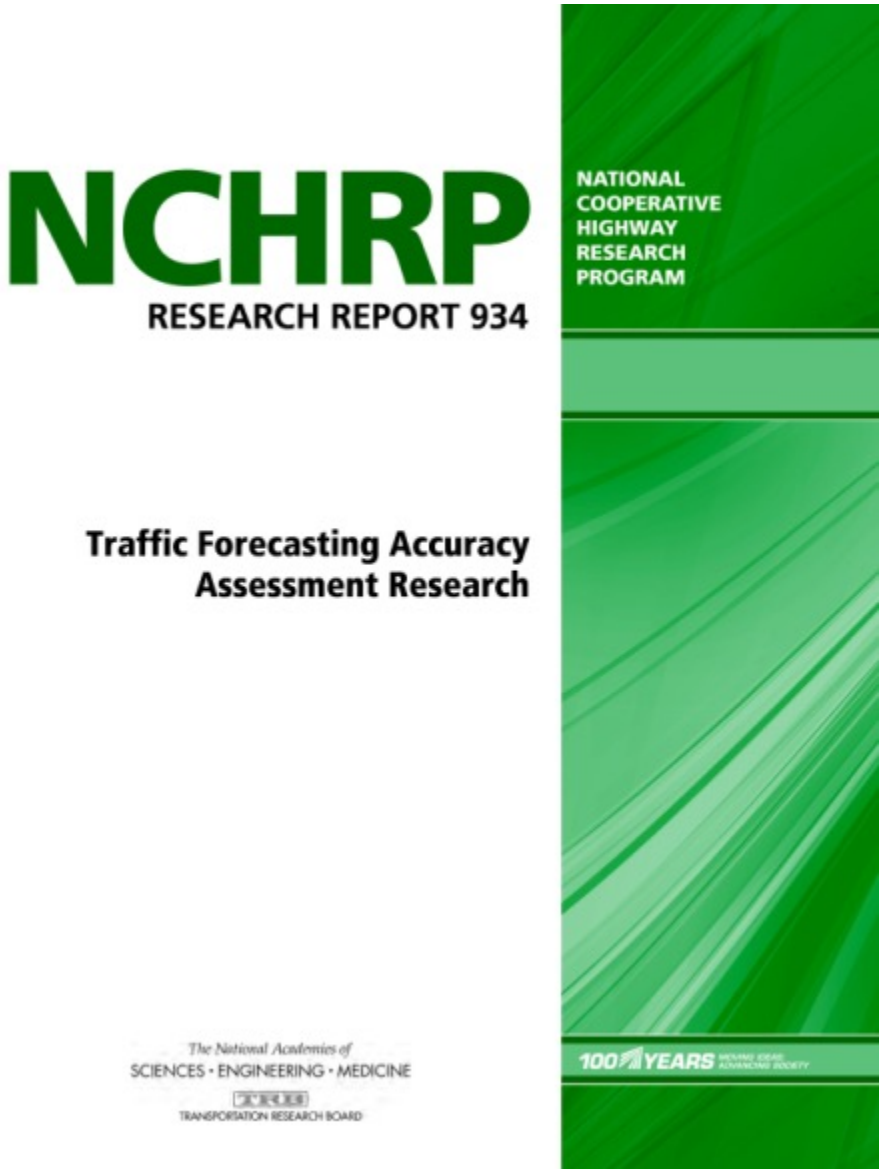
Suggestions for Practitioners

Use a range of forecasts to communicate uncertainty

Apply decision intervals to determine whether a forecasts at the high or low end of the range would change an investment decision

Systematically monitor traffic forecast accuracy and use the data to better estimate uncertainty

Further Readings

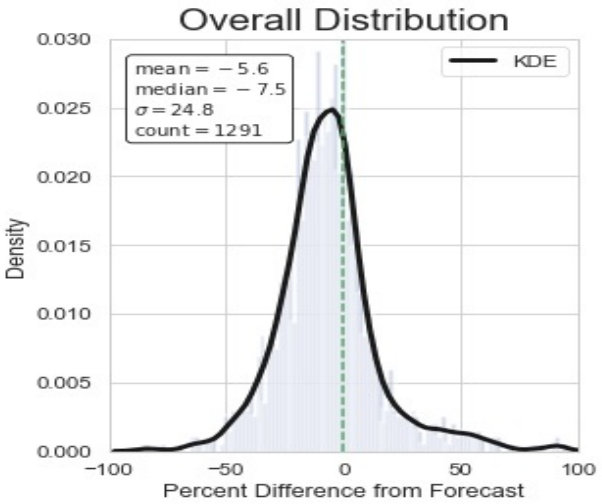


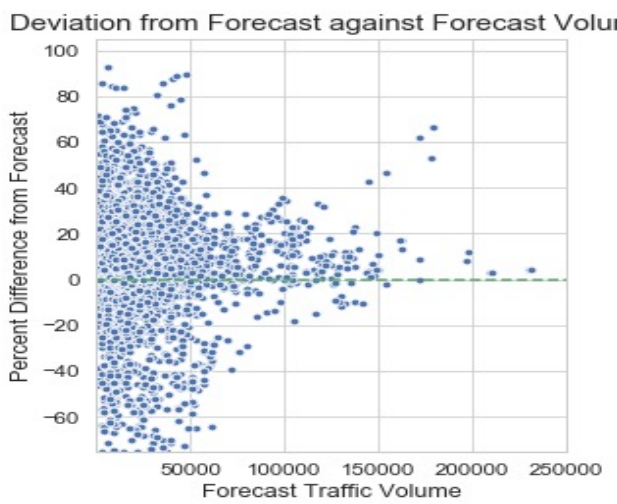
- **You can download a spreadsheet that implements these models from the NCHRP 934 website:**

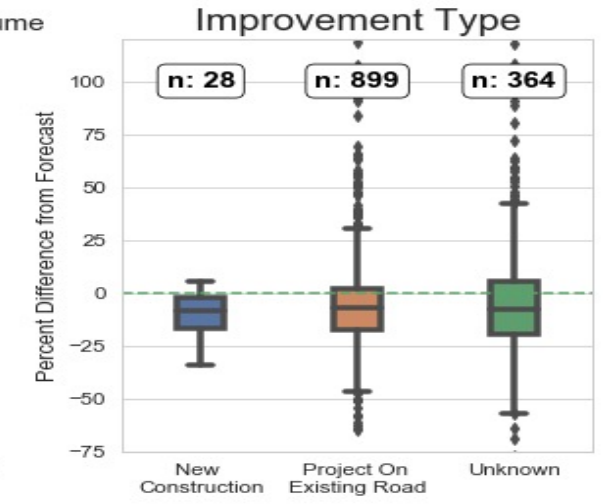
https://www.nap.edu/catalog/25637/traffic-forecasting-accuracy-assessment-research

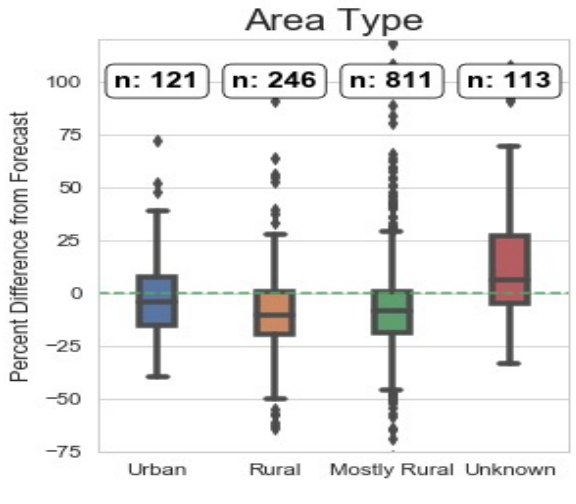
- NCHRP Research Report 934- Guidance Document
- Hoque, Jawad Mahmud, Gregory D. Erhardt, David Schmitt, Mei Chen, Ankita Chaudhary, Martin Wachs, and Reginald Souleyrette. 2021. "The Changing Accuracy of Traffic Forecasts." Transportation. https://doi.org/10.1007/s11116-021-10182-8
- Hoque, Jawad Mahmud, Gregory D. Erhardt, David Schmitt, Mei Chen, and Martin Wachs. 2021. "Estimating the Uncertainty of Traffic Forecasts from Their Historical Accuracy." Transportation Research Part A: Policy and Practice. https://doi.org/10.1016/j.tra.2021.03.015

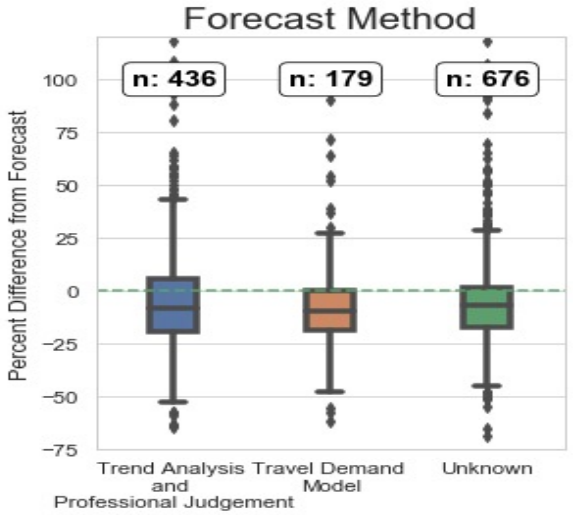
Traffic Forecast Accuracy

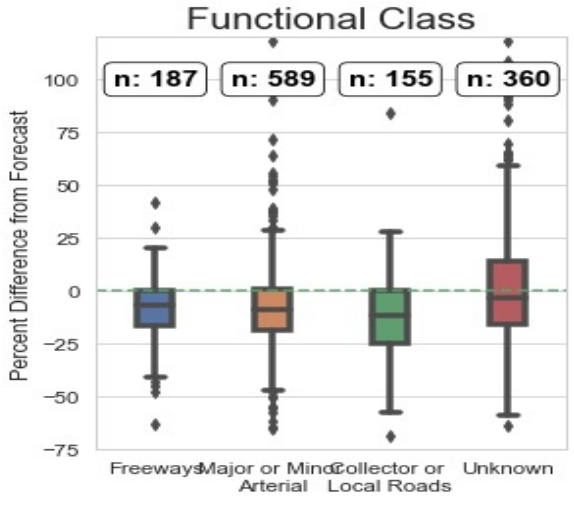


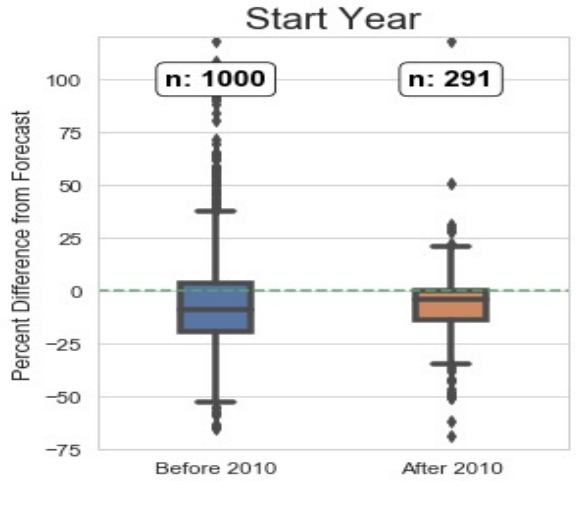


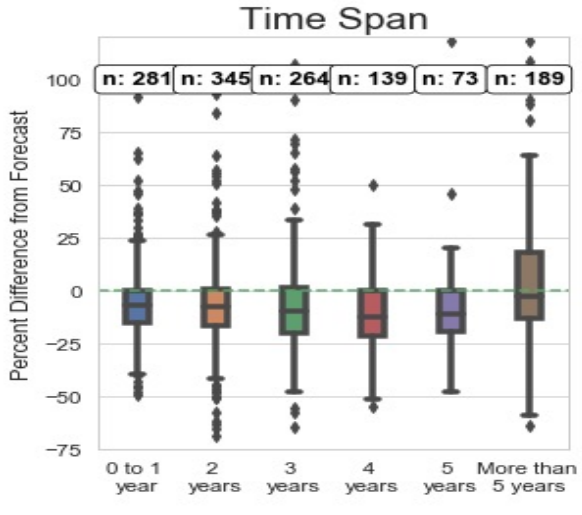












Source: Hoque, Jawad Mahmud, Gregory D. Erhardt, David Schmitt, Mei Chen, and Martin Wachs. 2021. "Estimating the Uncertainty of Traffic Forecasts from Their Historical Accuracy." Transportation Research Part A: Policy and Practice. https://doi.org/10.1016/j.tra.2021.03.015

Model Estimation Results- Equation Form

$$y_{5,i} = -182.26 + 0.705 * \hat{y}_i + 0.024 * \max(30{,}000 - \hat{y}_i, 0) + 0.006 * \text{TimeSpan} * \hat{y}_i - 0.006 * \text{UnemploymentRate} * \hat{y}_i + 1 * \text{Freeway} * \hat{y}_i - 0.15 * \text{Arterial} * \hat{y}_i - 0.212 * \text{Collector} * \hat{y}_i + 1 * \text{Existing Road} * \hat{y}_i + 0.093 * \text{New Road} * \hat{y}_i + 1 * \text{Trend Analysis} * \hat{y}_i + 0.068 * \text{Travel Model} * \hat{y}_i + 1 * \text{After 2010} * \hat{y}_i - 0.007 * \text{Before 2010} * \hat{y}_i + \varepsilon_{P,i}$$

$$y_{50,i} = 255.55 + 0.891 * \hat{y}_i - 0.004 * \max(30{,}000 - \hat{y}_i, 0) + 0.008 * \text{Time Span} * \hat{y}_i + 0.002 * \text{Unemployment Rate} * \hat{y}_i + 1 * \text{Freeway} * \hat{y}_i - 0.062 * \text{Arterial} * \hat{y}_i - 0.126 * \text{Collector} * \hat{y}_i + 1 * \text{Existing Road} * \hat{y}_i - 0.008 * \text{New Road} * \hat{y}_i + 1 * \text{Trend Analysis} * \hat{y}_i - 0.008 * \text{Travel Model} * \hat{y}_i + 1 * \text{Forecast Produced After 2010} * \hat{y}_i + 0.0002 * \text{Forecast Produced Before 2010} * \hat{y}_i + \varepsilon_{P,i}$$

$$y_{95,i} = 976.78 + 1.254 * \hat{y}_i - 0.413 * \max(30{,}000 - \hat{y}_i, 0) + 0.02 * \text{TimeSpan} * \hat{y}_i + 0.01 * \text{Unemployment Rate} * \hat{y}_i + 1 * \text{Freeway} * \hat{y}_i - 0.116 * \text{Arterial} * \hat{y}_i - 0.321 * \text{Collector} * \hat{y}_i + 1 * \text{Existing Road} * \hat{y}_i - 0.09 * \text{New Road} * \hat{y}_i + 1 * \text{Trend Analysis} * \hat{y}_i - 0.101 * \text{Travel Model} * \hat{y}_i + 1 * \text{After 2010} * \hat{y}_i + 0.003 * \text{Before 2010} * \hat{y}_i + \varepsilon_{P,i}$$