Improving Day Ahead Electricity Load Forecasts with Google Trends

Cameron Mulder
James Woods

July 29, 2014

Abstract

Modern short term load forecasting has grown in analytically complexity and sophistication. Day ahead forecasts now commonly use neural nets, Monte Carlo simulations and a wealth of historical data. What they have not done is fully captured the sentiment and intentions of the people using the electricity. This paper introduces Google Trend data, a summary of Google searches, as a way of capturing this sentiment and refining forecasts. We show with drop all forward cross validation that this amendment decreases forecast uncertainty by approximately 5% when compared to a statistically adjusted forecast and by over 50% when compared to raw forecasts.

1 Introduction

1. Intro to short term load forecasting.
2. Why crowd sourced, non technical, information could be useful.
3. Google trends is the summation of Google searches.
4. Outline of paper

2 Data Sources

2.1 PJM Load Forecasts and Actuals

```
> HourModelForecastCheck<-function(hour){
+   Hour<-formatC(hour, width=2, flag="0")
+   as.formula(paste("HE",Hour,"~ F",Hour,sep='' ))}
> SAResults<-lapply(1:24, FUN = function(x)  lm(HourModelForecastCheck(x), data=WTrends))
>
```

```
> library(ggplot2)
> ggplot(WTrends)+
+ geom_point(aes(F01,HE01,shape='a'))+geom_point(aes(F02,HE02,shape="b"))+
+ geom_point(aes(F03,HE03,shape='c'))+geom_point(aes(F04,HE04,shape="d"))+
+ geom_point(aes(F05,HE05,shape='e'))+geom_point(aes(F06,HE06,shape="f"))+
+ geom_abline(intercept = 0, slope = 1)
```

1. Data sources.
2. Documentation of forecasting.
3. Forecast bias

Figure 1: Confidence Intervals for Intercept (95%)

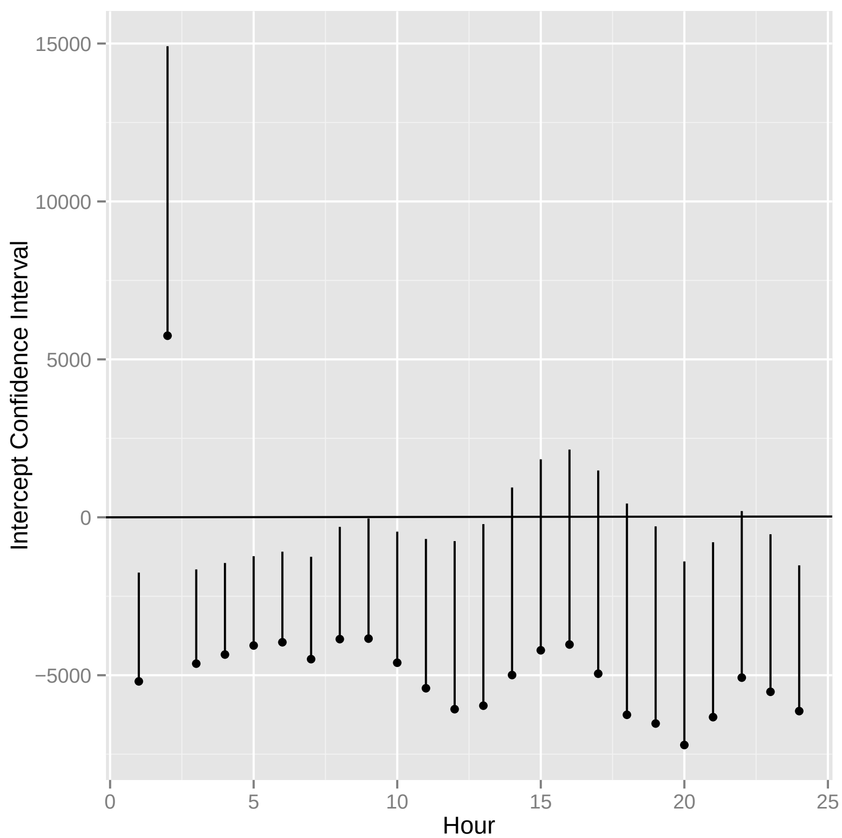
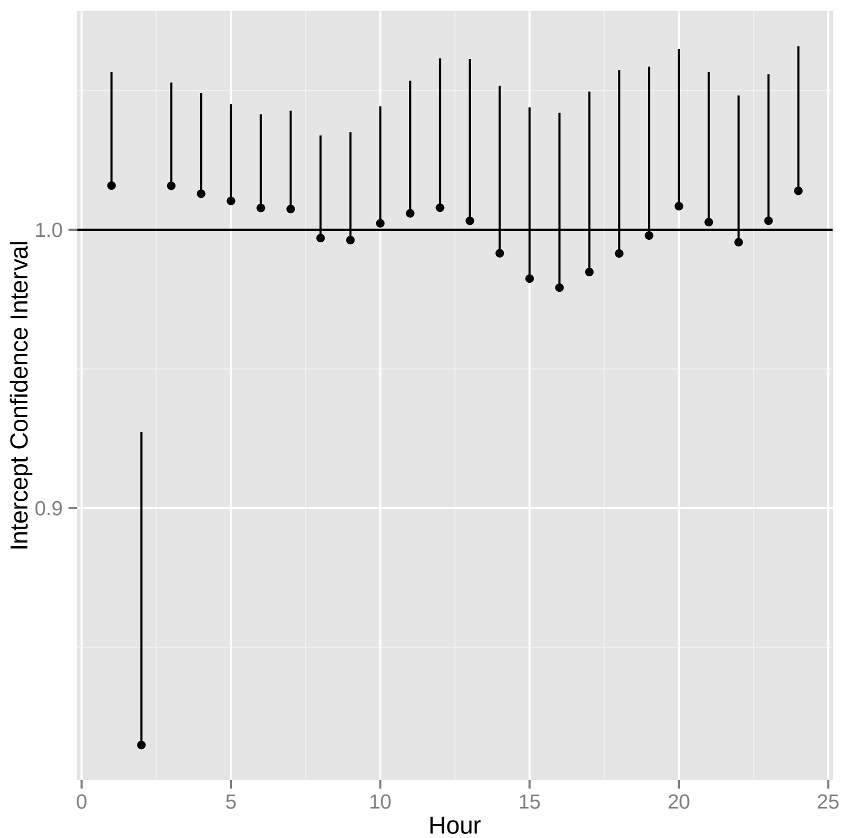


Figure 2: Confidence Intervals for Co-Movement (95%)



4. Statistically adjusted forecasts.

5. Note that almost all hours are biased and that co-movements are good for peak hours

2.2 Google Trends

1. Where to get the data

2. Limitations

3. Forming a population weighted index.

4. Other common searches that will be used as counter examples.

3 Post Forecast Addition of Google Trends Data

1. Simple hourly models with Trends.

2. Gross comparison with actual forecast and statistically adjusted forecasts.

3. Why this is insufficient.

3.1 Drop Forward Cross-validation

1. Cross validation concepts.

2. Why drop forward cross validation is the right concept.

3. Comparison of drop forward statistically adjusted and Trends adjusted with gross comparisons.

4. Reiteration that comparison with raw forecasts is a slam dunk.

3.2 Counter-factual Test with Other Common Google Searches

1. Comparison with: news, recipe, traffic, gas.

2. Note that some of them kinda work.

4 Summary and Conclusions