CSSS 504 Project Proposal

Giles Goetz, Robin Gold, Peter Schmiedeskamp

14 February 2014

Project Description

This project proposes to develop a statistical model for predicting the volume of bicycle traffic on the Fremont bridge in a given day, by controlling for factors such as weather, season, and day of the week. Observed bicycle counts have been collected on the bridge since late 2012 by an automated bicycle detection system(**City-of-Seattle:aa**).

A recent report prepared for the Puget Sound Regional Council (PSRC) reviews various factors influencing bicycle counts, including built-environment characteristics and time of day factors(**Bassok:2011aa**). This report corroborated findings by **City-of-Vancouver:1999aa**; **Niemeier:1996aa**; **Parkin:2008aa** of a relationship between bicycle counts, temperature and precipitation. However, it was not able to control for seasonal variation due to the lack of a sufficiently long-term dataset. A recommendation for further study was therefore for public agencies and researchers to consider installing automated counters that could be used to control for season.

In October, 2012 the City of Seattle installed automated bicycle counters on the Fremont Bridge, a major bicycling route in Seattle. This project will therefore be able to take advantage of a full calendar year's worth of bicycle count data to draw conclusions about seasonal variations and other potential long-term trends or factors affecting the number of cyclists on a given day.

Model Variables and Data Sources

The response variable is the daily aggregated bicycle count. The bicycle counts come from an automated detection system which uses an inductive loop installed in the sidewalk on each side of the bridge to register a passing bicycle. The system then reports an hourly aggregate count of bicycles crossing the bridge on each side, and this count data is archived and made publicly available on the City of Seattle website(**CoSwebsite**).

For this project, we chose to aggregate these hourly counts into daily counts in order to focus on the question of travel and commute mode choice in general, rather than peak-hour volumes or similar hourly factors. Additionally, we have divided the aggregated north-bound and south-bound counts by two, both because we anticipate that most trips will be two-way; and because bicycle traffic is two-way on both sides of the bridge, but the inductive loops make no distinction between bicycles traveling northbound or southbound.

Temperature and precipitation will be considered as predictors for the bicycle count, consistent with existing literature. Historic weather data is available through the Forecast.io public data service (**The-Dark-Sky-Company:aa**). Importantly, we will consider seasonality, which can be conceptualized at different levels of temporal resolution (e.g. 12 months versus 4 seasons). Other control variables include day of the week, holidays, and whether classes at the University of Washington are in session.

Additional Considerations

We have thus far identified two potential issues that may need to be accounted for in the regression model. The first is that the assumption of constant variance may be violated, by virtue of working with count data. The second issue is that there may be some dependence between weather observations (i.e. the weather one day is not fully independent of weather the next), leaving some potential for auto-correlation in the data set.