ESTIMATING DAILY BICYCLE COUNTS IN SEATTLE FROM SEASONAL AND WEATHER FACTORS

Peter Schmiedeskamp (Corresponding author)
University of Washington, Department of Urban Design and Planning
Box 355740, Seattle, WA 98195-5740
Tel: 206.289.0694
email: pschmied@uw.edu

Weiran Zhao
University of Washington, Department of Urban Design and Planning
Box 355740, Seattle, WA 98195-5740
Tel: 206.310.7065
email: weiran36@uw.edu

Word count: **6,952** (4,952 text & bibliography + 8 figures + 2 tables)

Submission Date: **1 August 2015**

ABSTRACT

This paper examines the relationship between several seasonal and weather factors and bicycle ridership based on two years of automated bicycle counts at a Seattle location. In so doing, we fitted a negative binomial model, and then estimated quantities of interest using counterfactual simulation. Our findings confirm the significance of season (+), temperature (+), precipitation (-), as well as holidays (-), day of the week (+ for Mon-Sat, relative to Sun.), and an overall trend (+). This paper improves on prior work by demonstrating the use of the negative binomial instead of a Poisson model, which is appropriate given the potential for overdispersion as observed in these data. In addition to validating the significance of factors identified from the literature, this paper contributes methodologically through its intuitive visualization of effect sizes to non-statistical audiences. We believe that the combination of model type and counterfactual simulation / visualization presented here reflect a reasonable compromise between model complexity and intepretability. Results such as these can aid policymakers and planners both in understanding bicycle travel demand elasticities, and to guide interventions aimed at increasing rates of bicycling. Finally, the methods presented here are fully reproducible and invite adaptation to other locations.

Keywords: Bicycle counts; negative binomial; season; weather; automated; counterfactual simulation; visualization

INTRODUCTION

Bicycling offers a variety of social benefits that include the health of those choosing to bicycle regularly, as well as, more broadly, a society-wide increase in mobility when people shift from congestion causing modes to bicycling. As such, bicycling is being promoted by policy makers and urban planners as an increasingly important component of urban transportation. In order to develop policy and infrastructural improvements to induce more bicycling, a robust understanding of factors related to bicycling is required.

This article contributes to that understanding by focusing on the substantive effects of weather and seasonality on the number of bicyclists observed on a given day. Specifically, we hypothesize that factors associated with inclement weather (esp. low temperature and precipitation) will be inversely correlated with bicycle counts. We similarly expect to see an additional decrease in counts associated with both physical and social / institutional dimensions of season. That is, in seasons where the days get shorter, we expect fewer cyclists due to a kind of decision making inertia. We also expect fewer cyclists outside the "school season" as we believe that a large number of bicycling trips are associated with commuting to schools that, unlike most other institutions have clear annual cycles.

In order to address these two primary research questions, we develop a statistical model estimated using two years of automated bicycle counts collected on Seattle's Fremont Street bridge, as well as historic weather and season data collected from other sources.

Using this model, we were able to identify and quantify substantial effects on bicycle counts associated with temperature, precipitation, hours of daylight, and school session status. Additionally, we estimated changes in bicycle numbers associated with the day of the week and holidays, as well as an overall trend toward increased bicycling.

LITERATURE REVIEW

Bicycling volume in cities is useful for practitioners and researchers to understand safety, travel behavior, and development impacts. Therefore the relationship between bicycle volume and various factors, with the goal to build a predictive model based on this relationship, has been of great interest to researchers over the last decade (e.g. Griswold, Medury, and Schneider 2011; Fields 2012; Niemeier 1996; Nosal and Miranda-Moreno 2014).

To address these questions, researchers have used a variety of approaches using different data sources. Studies employing survey or census data are more often used to explain influences such as demographic and socioeconomic factors on mode choice (Parkin, Wardman, and Page 2008; Helbich, Böcker, and Dijst 2014). In other studies, collected count data (from automated counters or manual counts) have more been used recently to track and analyze counts over longer periods of time (Griswold, Medury, and Schneider 2011; Nosal and Miranda-Moreno 2014). These counts-based approaches have largely followed a similar approach of first proposing a set of explanatory variables, fitting some form of regression model, and then interpreting / justifying the results visà-vis the guiding theory.

A literature review accompanying a recent report by Bassok et al. (2011) identified eleven primary indicators. These included time of day (Schwartz et al. 1999), season (Niemeier 1996), population and employment densities (McCahil 2008; Pinjari, Bhat, and Hensher 2009), land-use mix (Pinjari, Bhat, and Hensher 2009), bicycle facility type (Hunt and Abraham 2007), traffic volume (McDonald and NZ. 2007), rain and temperature (Niemeier 1996; Parkin, Wardman, and Page 2008), income (Turner et al. 1998), and age (Hunt and Abraham 2007). This section outlines the key points made in the literature that are relevant to some of most important variables.

Research has found the variability for counts has a positive association with high temperature and low precipitation. (Niemeier 1996; Parkin, Wardman, and Page 2008). Meanwhile, as suggested by Lewin (2011) and Thomas, Jaarsma, and Tutert (2009), the effects of precipitation and temperature on bicycle volumes are nonlinear. For example, bicycle traffic can decrease in both very cold and very hot weather as noted by Richardson (2000). Apart from the usual temperature and rain variables, Miranda-Moreno and Nosal (2011) finds humidity and additional precipitation variables including the presence of rain in the morning and/or during the previous three hours to be significant too. Other comparative studies are also available where bicycle counts are conducted in different cities, and different sensitives to weather are examined (Rose et al. 2011). As for longitudinal studies, Niemeier (1996) finds increased variability for counts conducted in the later months of the year. Jones et al. (2010) conclude that morning peak hours from 6 AM to 9 PM accounts for a consistent 95% of the total bicycle volumes by hourly count data. The simple linear regression model has been used in several applications (Jones and Buckland 2008; Jones et al. 2010). Other modeling approaches include Miranda-Moreno and Nosal (2011) which develops a count model and Thomas, Jaarsma, and Tutert (2009) which develops a time-series model. Niemeier (1996) also uses a Poisson model to statistically confirm many of the factors thought to influence cyclists. The work by Gallop, Tse, and Zhao (2012) adopts a similar time-series approach while incorporating an autoregressive integrated moving average (ARIMA) analysis. A summary of findings from the literature are presented in Table 1.

These literature together suggest an opportunity for further model development base around long-term automated counts utilizing appropriate statistical methods. How seasonal factors influence bicycle flow needs to be examined in data that last more than a year. One limitation present in much of the past literature is that few discuss goodness of fit of their modeling. Further a model that can better describe and forecast the bicycle count in longitudinal form is necessary to be developed. Models for count data with better estimation methods offer some promise.

TABLE 1 Summary of findings from literature review

Source	Variable(s) identified	Methods
Fields 2012	Average daily temperature; Total weekly precipitation	Identify patterns through scatter plots; No explicit model is established.
Gallop, Tse, and Zhao 2012	Temperature, Relative humidity, wind speed, visibility, fog, precipitation	Use ARIMA to account for serial correlation patterns
Griswold, Medury, and Schneider 2011	Nearby population and employment density, proximity to downtown/freeway, age, education level, income, etc.	Log linear ordinary least squares regression is used to estimate a bicycle count model
Helbich, Böcker, and Dijst 2014	Daily maiximum air temperature, daily average wind speed, daily precipitation	Place-specific associations of weather conditions are explored through geographically weighted logit models
Hunt and Abraham 2007	Descriptive variables indicating lane use, secured parking, level of experience, etc.	Logistic model of cycling-related choices
Jones et al. 2010	Length of bicycle network, employment density, population density	Standard ordinary least squares regression
Lewin 2011	Max temp, rain flag, snow flag, weekend flag, over 90 flag	Standard linear regression model
McCahil 2008	logarithmic choice measure, population density, worker density	A new space syntax theory is used to evaluate and predict the bicycle volume throughout a network
Miranda-Moreno and Nosal 2011	Temperature, percent humidity, rain presence, rain presence in prev. 3hrs, warm & humid, morning rain	Both log-linear model and negative binomial model are tested
Niemeier 1996	Morning flag, rain flag, high temp flag, location variable, season variable	A Poisson count model is assumed and fitted
Nosal and Miranda-Moreno 2014	Temperature, percent humidity, rain flag, rain prev. 3hrs, am rain, pm rain	The relationship is analyzed using a log-linear regression model
Parkin, Wardman, and Page 2008	Gender, car ownership, hilliness, off-road routes proportion	A logistic regression model of relevant socio-economic and physical variables is estimated.
Pinjari, Bhat, and Hensher 2009	Household density, employment density, fraction of commercial land area, demographic factors including proportion of population that are seniors and proportion of population by race	The model system takes the form of a joint mixed Multinomial LogitâĂŞMultiple Discrete-Continuous Extreme Value (MNL-MDCEV) structure
Rose et al. 2011	Temperature, rainfall, holiday flag, school season flag, day of the week	Weather and other effects examined using an aggregate model of daily ridership
Thomas, Jaarsma, and Tutert 2009	Temperature, sunshine, precipitation, wind force, cycle path use	A bi-level structure is developed with the upper level being a log-linear model and the lower level being a linear model

METHODOLOGY

In order to discern the relationship between bicycle counts and several identified weather, seasonal, and temporal factors, we developed a statistical model that attempts to predict daily bicycle counts from these other factors. This section describes the methods and procedures we used to collect and process the estimation dataset, the rationale for our selection of variables, our chosen model type, and our model estimation procedures.

Data collection, processing, description

The data that we attempt to predict were collected at Seattle's Fremont Bridge and cover a period of two years spanning from October 31, 2012 to October 30, 2014. The Fremont Bridge captures a substantial amount of bicycle traffic due to its status as one of only five facilities that carry bicyclists across the canal separating the northern and southern halves of Seattle.

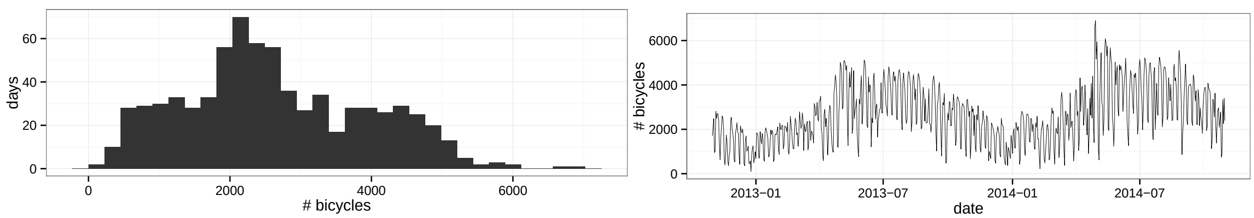
Bicycle counts were collected at this location continuously by the City of Seattle using an in-sidewalk counter manufactured by EcoCounter. When a bicycle passes over an induction loop embedded in the sidewalk on either side of the Fremont Bridge, the counter registers the bicycle. Bicyclists may legally choose to ride in the roadway instead of the sidewalk, and would thus not be detected by the counter. However, we believe these crossings are rare at this location due to the design of the facility, which directs bicyclists to enter the sidewalk, and from our own experience riding and observing other riders. Counts are aggregated into 15 minute intervals by the City of Seattle, and are made available to the public via the City of Seattle's data portal (City of Seattle 2013, 2014).

Weather data are collected by a variety of sources and are aggregated by Forecast.io. These data are available through the company's web services API (The Dark Sky Company 2013). Historical daily summaries are available for a range of weather variables including several specifically important to our model such as precipitation, daily minimum and maximum temperatures, sunrise, and sunset.

We downloaded and processed these data programmatically using the R programming language along with several add-on packages (Grolemund and Wickham 2011; Wickham 2011; Couture-Beil 2014; Lang 2014; R Core Team 2014).[1] Bicycle counts were aggregated by day, and then joined to weather data by date. In addition to the variables collected from these two sources, we were also interested in controlling for holidays and whether or not the nearby University of Washington was in session. These data were collected and coded manually from the University of Washington's historic academic calendars.

Figure 1 provides a visual summary of the processed counts data. Some apparent outliers are visible at the rightmost portion of the histogram. The two highest counts occurred the Monday and Tuesday preceding the beginning of National Bike to Work Month. And the third highest count occurred on National Bike to Work Day.

[1]In addition, the code used to run the analysis, as well as the final dataset are available as a github repository: `https://github.com/pschmied/bikecounts-mle`



(a) Frequency of observed counts (b) Timeseries of counts

FIGURE 1 Descriptive visualization of bicycle counts dataset.

Variable selection

Out of the set of independent variables retrieved from Forecast.io and the University of Washington, we selected a subset that we felt best reflected our specific research questions. The final set of variables included in our model include those presented and summarized in Table 2. Further explanation and justification of these variable choices follow.

TABLE 2 Variables included in model specification

Variable	Description
Count*	Number of bicycles per day
Daylight	Length of daylight in hours
UW	The University of Washington was in session (TRUE/FALSE)
Holiday	The day was a holiday as recognized by UW (TRUE/FALSE)
Max temp	Maximum temperature in Fahrenheit for the day
Max temp^2	`Max temp` squared
Max precip	Daily maximum inches of precipitation in any hour
Sat–Fri	Day of the week dummy variable (TRUE/FALSE, relative to Sunday)
Day #	Sequentially numbered day of study

* Dependent variable

Daylight hours (defined as $sunset - sunrise$) and the University of Washington in-session status were selected to represent seasonality. Daylight hours was chosen instead of a calendar-based categorization of season in part because Seattle's Pacific Maritime Climate differs substantially from traditional notions of Spring, Summer, Autumn, and Winter. Daylight hours also is measured as a continuous value at a finer temporal resolution of one day. Finally, daylight hours adjusts according to latitude, which may make this model estimation procedure and specification more transferable to other sites in the future, perhaps by interacting latitude with daylight hours.

We deemed the University of Washington variable important in part because of the Fremont Bridge's proximity and connection via the Burke Gilman Trail to the University of Washington. We also felt that this variable was a suitable proxy for the "school season," which more broadly captures whether or not other local schools are in session. The academic calendars of the various local schools do not align perfectly, however they still overlap substantially with the University of Washington, which is itself the largest educational institution in the region.

Inclusion of the holiday variable was an attempt to account for some low outlier counts. Upon inspection of the dataset, Christmas and Thanksgiving in particular had very low counts of bicycles relative to the days preceding and following. Relatedly, but not accounted for by any variable in our model, are some of the high outlier counts. Upon inspection, some of the highest counts were observed on National Bike to Work Day and on the day of the Fremont Solstice Parade, which typically draws large numbers of bicyclists as participants and spectators. The omission of such a variable is justified based on the few occurrences of high outlier counts, and our desire for this model to only include variables that could be collected or straightforwardly adapted to other locations.

Daily maximum temperature, measured in Fahrenheit, was chosen to represent temperature (rather than, for example, substituting or adding daily minimum temperature) in part to retain simplicity in the model, in part because there is relatively little daily temperature variation in Seattle due to the moderating effect of large water bodies, and in part because maximum temperature better reflects the conditions during daylight hours when most bicycle trips would occur. This simplification may not be warranted for other locations that experience greater temperature variation than Seattle. In addition to daily maximum temperature, we also included a daily maximum temperature squared term in order to identify the potential for a leveling off of counts in very high temperatures.

Maximum precipitation, which measures the maximum inches of precipitation that occurred in any hour throughout the day, was chosen rather than average precipitation based on the notion that bicyclists might make travel decisions based on a likely worst case scenario. This assumption is slightly more problematic than our assumptions about temperature, in that we do expect bicyclists to be at least somewhat sensitive to average conditions or conditions observed at their time of departure. As in the case of temperature, this simplification would be less justifiable in locations that experience greater daily variation in precipitation or in locations that have a predictable pattern of precipitation during certain hours.

Day of the week was added due to its presence in the literature, as well as an apparent weekly pattern revealed visually by zooming into the timeseries plot. These data were coded as a set of Boolean dummy variables, excluding Sunday as the reference category.

The final variable, the day number, was included so that we could test for a linear trend in bicycling volumes. We created this variable by sequentially numbering (1–720) the observed counts by day during the study period.

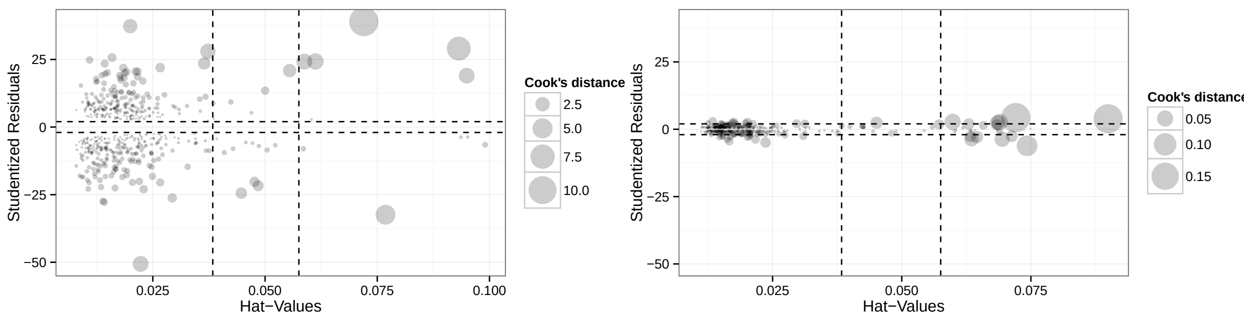
Model estimation and goodness of fit

A natural model choice for count data is the Poisson model, however we believed that our count data were overdispersed. Overdispersion, or contagion between events, violates the mean-variance equivalence assumption of the Poisson model. While overdispersion would not impact our parameter estimates, it would result in overly optimistic margins of error. In order to account for and estimate the amount of overdispersion present in our data, we chose a negative binomial model type. Because there were no days observed with zero bicycle counts, we did not need to resort to zero-inflated models as is often necessary when dealing with count data.

We fit the model in R using the `glm.nb` function from the MASS package (Venables and Ripley 2002). For comparison, we also estimated a Poisson model with an analogous specification in R using the `glm` function. A much lower AIC and BIC in the case of the negative binomial model confirmed that it was a better fit than the Poisson. We then tested for overdispersion using the `odTest` function from the Pscl package (Jackman 2014). The highly significant chi-squared test statistic provided strong evidence that overdispersion was present in the data.

Additionally, we compared the residuals plots for both the Poisson and negative binomial models. As seen in figure 2, the residual variance is much lower in the negative binomial model, further confirming the choice of a negative binomial for these counts.

As a final model comparison, we visually assessed the overall fit of our negative binomial model by plotting actual versus predicted values as shown in figure 3. This fit appears to be generally good, though very high count days are predicted somewhat less accurately.



(a) Poisson (b) Negative binomial

FIGURE 2 Residual variance of Poisson vs. negative binomial model.

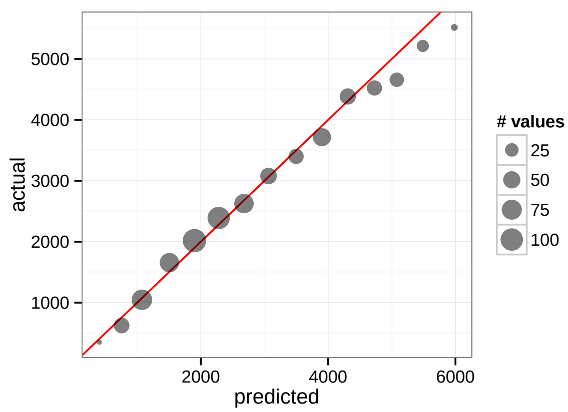


FIGURE 3 Actual versus predicted values as fit by negative binomial model.

In order to provide results that are more readily interpretable by non-statisticians, we used counterfactual simulation to isolate individual terms from the model that correspond to our research questions. In so doing, we simulated various quantities of interest including point estimates and confidence intervals, and then plotted them for visual inspection. Counterfactual simulations were performed with a modified version of the Simcf R package, and visualized with ggplot2 (Adolph 2014; Schmiedeskamp 2014; Wickham 2009). Results of these simulations are presented in the following section.

RESULTS

This section presents the results from the statistical model and accompanying counterfactual simulations as described in the preceding section. Each of the main research questions of seasonality, weather, and general trend are addressed here. In addition, additional added control variables such as day of the week, holidays, and linear trend are presented.

With the exception of the `Max temp`^{2} term, each of the coefficients in our model was statistically significant at the $p < 0.05$ level. Further, with the exception of the `Max temp`^{2} and `Saturday` coefficients , all coefficients were significant at the $p < 0.001$ level. While this is instructive in confirming the significance of these factors, the remainder of this section focuses on presenting the substantive effect of each variable.

Seasonality

As discussed previously, we considered two variables to address the question of seasonality: the first being the number of daylight hours, and the second being whether or not the University of Washington (proxying more generally for other educational institutions) was in session.

Figure 4 shows that we see a substantial increase in bicycle volumes when the University of Washington was in session. With all other factors held constant we see that, on days when the university is in session, we would expect an average of approximately 367 additional bicycles observed. Similarly, we see a roughly linear increase in bicycles with increased day length.

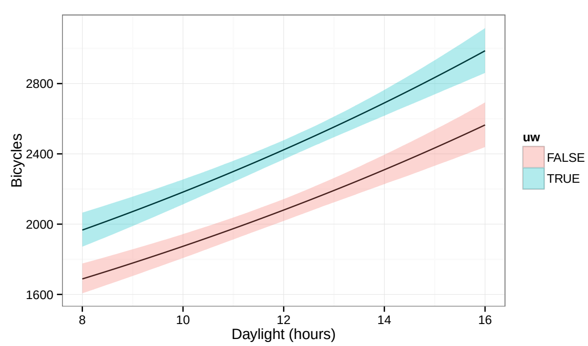


FIGURE 4 Effect of daylight hours and University of Washington in-session status on bicycle counts, with shaded 95% confidence regions.

Weather

For our examination of seasonality, our model represents weather using three variables. In this case, these are precipitation (measured as the maximum amount of precipitation falling in any hour of that day), maximum temperature (measured daily, in Fahrenheit), and maximum temperature squared.

The effect of precipitation is shown in figure 5. From this, we can see a clear inverse relationship between precipitation and bicycle counts. The rate of decrease in bicycles appears to begin somewhat steeply, and then begins to slow slightly at higher amounts of precipitation. This suggests that people are generally more sensitive to the presence of precipitation than the intensity.

Temperature, in contrast to precipitation, has a clearly positive association with increased numbers of bicyclists. Temperature squared was not significant in our model at the $p < 0.05$ level. We were somewhat surprised to not see evidence for a leveling off in counts at very high temperatures—likely due to Seattle's moderate summer climate, and the lack of extremely warm days in our dataset.

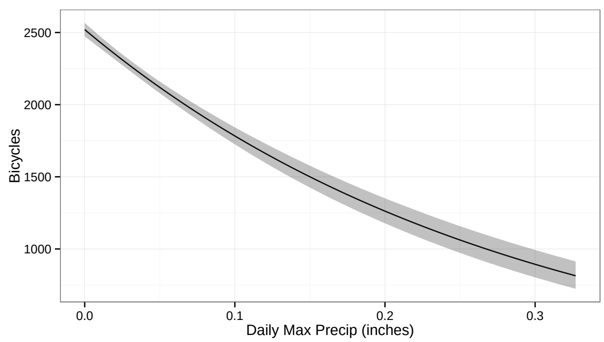


FIGURE 5 Effect of precipitation on counts, with shaded 95% confidence region.

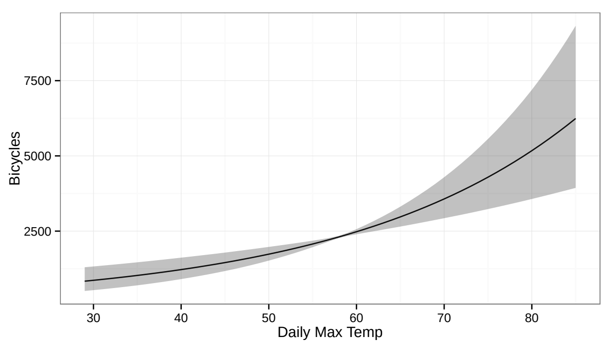


FIGURE 6 Effect of temperature on counts, with shaded 95% confidence region.

General trend in bicycle counts

Our results (shown in figure 7) confirm the presence of a general trend toward increased numbers of bicycles at this location. With all else held constant, we would expect to see roughly 300 more bicycles on days at the end of our study period than at the beginning. This is consistent with figures reported elsewhere that suggest that bicycling is increasing in a number of cities including Seattle (League of American Bicyclists 2014).

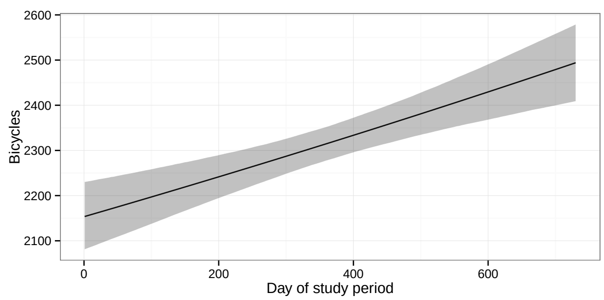


FIGURE 7 General trend in bicycling counts, all other factors held constant, with 95% confidence region.

Day of week variation

As discussed previously, day of the week was included due to the weekly variation in bicycle counts apparent in timeseries plots. Figure 8 shows several interesting aspects of these results. First, we see much higher numbers of bicyclists on weekdays than on weekends. This strongly suggests that the majority of the bicycle traffic at this location is for commuter purposes.

Comparing between weekday results, we see that most days are roughly the same, with the some drop-off toward Thursday and Friday. This decrease toward the end of the work week might be attributable to individuals either working non-traditional schedules or perhaps people adjusting their travel mode choice in order to accommodate social engagements.

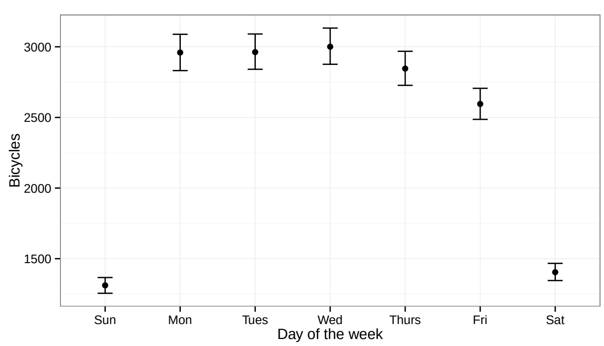


FIGURE 8 Variation in counts throughout the week, with 95% confidence bars.

DISCUSSION

This research set out to better understand the relationship between bicycle counts and weather and season. However, weather and season are themselves comprised by a number of constituent elements. We believe this research has chosen defensible representations of weather and season. However, we readily concede that there exist other reasonable representations of season and weather.

In particular, others might be more interested in building a model that places more emphasis on predictive power. Those individuals might choose to add a number of additional variables or consider more complex interactions between terms. By contrast, in this model formulation, we present a relatively simple model with comparatively fewer terms. This was due to our interest in a conceptual understanding of the influence of weather and season.

Even within our model, as noted in the methodology section, we considered alternative specifications, especially those containing different measures of temperature and precipitation. As an example, in choosing to represent temperature as the daily maximum, we make some assumptions about people and their mode choice decision making processes. People cannot know the daily maximum temperature in advance of their morning commute, however the measured daily maximum temperature is likely to be similar to temperatures reported in temperature forecasts. A superior measure might be a consensus of forecasts available in the morning of each day. Similarly, if Seattle's daily high and daily low temperatures did not have such low variation, a daily maximum temperature might be less important to a winter-time bicyclist than daily minimum.

One variable missing from our analysis that we would have liked to include is cloud cover. This variable would have allowed us to account for the possibility of brilliant, sunny, but cold winter days, as well as warm, dry, but gloomy summer days. Unfortunately, the dataset downloaded from Forecast.io included a great deal of missing values for this variable, and thus future work might include identification of a more suitable source of these data.

Another limitation of this research comes from our sample of bicycle counts. While we do have continuous counts spanning a full two years, these counts were taken at just one point location. We believe the Fremont bridge is somewhat representative of a high-volume bicycle facility in Seattle. However we might, for example, see a decreased effect size for University of Washington session status if we looked at counts further from the university. The City of Seattle has begun collecting and releasing data from other counters, which could be included in a future version of this analysis.

Finally, as discussed in the methodology section, we believe the choice of the negative binomial model to be a better choice than the Poisson, which had been used in some previous studies. A limitation in this model, however, is in not accounting for timeseries autocorrelation. We think it highly likely that the decision to bicycle on one day might influence the decision to bicycle on the next. A future direction would be to consider models, such as ARIMA, that more explicitly account for time, though that decision would represent a potentially undesirable increase in model complexity.

CONCLUSIONS

This research set out to help clarify the relationship between weather and seasonal factors and bicycle counts in Seattle. In order to achieve this, we developed a negative binomial model to predict bicycle counts based on temperature, precipitation, day length, university in-session status, day of the week, and a general linear time trend.

For each term included in our model except temperature squared, we found statistically significant effects. More importantly, through the use of counterfactual simulation, we estimated what we deem to be substantial effect sizes associated with each predictor variable.

This article contributes to the existing literature by demonstrating the use of an appropriately chosen negative binomial model for bicycle counts. It provides an additional methodological contribution in illustrating counterfactual simulation and visualization to create more compelling and intuitive results summaries for non-statisticians. Finally, the results presented here were generated from data collected over a relatively long study period of two years.

While control of the weather and seasons are admittedly beyond the scope of policy makers, this research does suggest that planners and policy makers may want to develop strategies that help mitigate the impacts of the natural environment during the winter months. In other words, the delta between warm dry days and cold wet days should be treated as the opportunity frontier. Future research could focus on determining what, if any, programmatic or built interventions could ameliorate unfavorable cold- and wet-weather bicycling conditions.

ACKNOWLEDGMENT

We acknowledge Profs. Elena Erosheva and Christopher Adolph for their guidance on the development of this topic. We also acknowledge Giles Goetz and Robin Gold for their involvement on a prior incarnation of this work.

REFERENCES

Adolph, Christopher. 2014. *simcf: Counterfactuals and confidence intervals for estimated models.* R package version 0.3.0.

Bassok, Alon, Noa Ginger, Andy Hong, John Murphy, Danielle Rose, Peter Schmiedeskamp, Amanda Snypp, and Eiji Torikai. 2011. *Bicycle Planning, Best Practices, and Count Methodology.* Technical report. Seattle: Puget Sound Regional Council and University of Washington Department of Urban Design and Planning.

City of Seattle. 2013. "Fremont Bridge Bike Counter." `http://www.seattle.gov/transportation/bikecounter_fremont.htm`.

———. 2014. "Open Data." `https://data.seattle.gov`.

Couture-Beil, Alex. 2014. *rjson: JSON for R.* R package version 0.2.15. `http://CRAN.R-project.org/package=rjson`.

Fields, Billy. 2012. *Active Transportation Measurement and Benchmarking Development: New Orleans State of Active Transportation Report 2010.* Paper 4. UNOTI Publications. `http://scholarworks.uno.edu/unoti_pubs/4`.

Gallop, Christopher, Cindy Tse, and Jinhua Zhao. 2012. "A Seasonal Autoregressive Model of Vancouver Bicycle Traffic Using Weather Variables." *Transportation Research Board 2012 Annual Meeting.*

Griswold, Julia B., Aditya Medury, and Robert J. Schneider. 2011. *Pilot Models for Estimating Bicycle Intersection Volumes.* Institute of Transportation Studies, Research Reports, Working Papers, Proceedings. Institute of Transportation Studies, UC Berkeley. `http://EconPapers.repec.org/RePEc:cdl:itsrrp:qt380855q6`.

Grolemund, Garrett, and Hadley Wickham. 2011. "Dates and Times Made Easy with lubridate." *Journal of Statistical Software* 40 (3): 1–25. `http://www.jstatsoft.org/v40/i03/`.

Helbich, Marco, Lars Böcker, and Martin Dijst. 2014. "Geographic heterogeneity in cycling under various weather conditions: evidence from Greater Rotterdam." *Journal of Transport Geography* 38:38–47. `http://www.sciencedirect.com/science/article/pii/S0966692314000957`.

Hunt, J.D., and J.E. Abraham. 2007. "Influences on bicycle use" [in English]. *Transportation* 34 (4): 453–470. `http://dx.doi.org/10.1007/s11116-006-9109-1`.

Jackman, Simon. 2014. *pscl: Classes and Methods for R Developed in the Political Science Computational Laboratory, Stanford University.* R package version 1.4.6. Stanford, California: Department of Political Science, Stanford University. `http://pscl.stanford.edu/`.

Jones, Michael G., Sherry Ryan, Jennifer Donlon, Lauren Ledbetter, David R. Ragland, and Lindsay Arnold. 2010. *Seamless Travel: Measuring Bicycle and Pedestrian Activity in San Diego County and its Relationship to Land Use, Transportation, Safety, and Facility Type.* Technical report. Berkeley: UC Berkeley Safe Transportation Research & Education Center, Caltrans Task Order 6117, June.

Jones, Michael, and Lauren Buckland. 2008. *Estimating Bicycle and Pedestrian Demand in San Diego.* Technical report. Berkeley: UC Berkeley Safe Transportation Research & Education Center, Caltrans Task Order 6117.

Lang, Duncan Temple. 2014. *RCurl: General network (HTTP/FTP/...) client interface for R.* R package version 1.95-4.3. `http://CRAN.R-project.org/package=RCurl`.

League of American Bicyclists. 2014. "Updated: Bike commute data released." `http://bikeleague.org/content/updated-bike-commute-data-released`.

Lewin, Amy. 2011. "Temporal and Weather Impacts on Bicycle Volumes." *Transportation Research Board 90th Annual Meeting.*

McCahil, Chris. 2008. "The Applicability of Space Syntax to Bicycle Facility Planning." *Transportation Research Record: Journal of the Transportation Research Board* 2074 (1): 46–51. `http://www.sciencedirect.com/science/article/pii/S0191261509000198`.

McDonald, A.A., and New Zealand. Land Transport NZ. 2007. *Estimating Demand for New Cycling Facilities in New Zealand.* Land Transport New Zealand Research Report Series. New Zealand Government - NZ Transport Agency. `http://books.google.com/books?id=iNiGNAAACAAJ`.

Miranda-Moreno, Luis F., and Thomas Nosal. 2011. "Weather Or Not to Cycle; Whether Or Not Cyclist Ridership has Grown: A Look at Weather's Impact on Cycling Facilities and Temporal Trends in an Urban Environment." *Transportation Research Board 2011 Annual Meeting.*

Niemeier, Debbie A. 1996. "Longitudinal Analysis of Bicycle Count Variability: Results and Modeling Implications." *Journal of Transportation Engineering* 122 (3): 200–206.

Nosal, Thomas, and Luis F. Miranda-Moreno. 2014. "The effect of weather on the use of North American bicycle facilities: A multi-city analysis using automatic counts." *Transportation Research Part A: Policy and Practice* 66:213–225. http://www.sciencedirect.com/science/article/pii/S0965856414000998.

Parkin, John, Mark Wardman, and Matthew Page. 2008. "Estimation of the Determinants of Bicycle Mode Share for the Journey to Work Using Census Data." *Transportation* 35 (1): 93–109.

Pinjari, Abdul Rawoof, Chandra R. Bhat, and David A. Hensher. 2009. "Residential self-selection effects in an activity time-use behavior model." *Transportation Research Part B: Methodological* 43 (7): 729–748. http://www.sciencedirect.com/science/article/pii/S0191261509000198.

R Core Team. 2014. *R: A Language and Environment for Statistical Computing*. Vienna, Austria: R Foundation for Statistical Computing. http://www.R-project.org/.

Richardson, A. J. 2000. "Seasonal and Weather Impacts on Urban Cycling Trips." *TUTI Report.*

Rose, G, F Ahmed, M Figliozzi, and C Jakob. 2011. "Quantifying and comparing the effects of weather on bicycle demand in Melbourne (Australia) and Portland (USA)." *TRB 90th Annual Meeting:*

Schmiedeskamp, Peter. 2014. "Fork of Simcf." https://github.com/pschmied/tile-simcf.

Schwartz, W. L., Cambridge Systematics., Bicycle Federation of America., and Turner-Fairbank Highway Research Center. 1999. *Guidebook on methods to estimate non-motorized travel [microform] : supporting documentation / [authors, W.L. Schwartz ... et al.]* [in English]. 1 v. (various pagings) : U.S. Dept. of Transportation, Federal Highway Administration, Research, Development, / Technology, Turner-Fairbank Highway Research Center ; Available to the public through the National Technical Information Service McLean, VA (6300 Georgetown Pike, McLean 22101-2296) : [Springfield, Va.

The Dark Sky Company. 2013. "Forecast API Docs." https://developer.forecast.io/docs/v2.

Thomas, T, R. Jaarsma, and B. Tutert. 2009. "Temporal Variations of Bicycle Demand in the Netherlands: Influence of Weather on Cycling." *CD Proceedings of the 88th Annual Meeting of the Transportation Research Board.*

Turner, S., G. Shunk, A. Hottenstein, Texas. Dept. of Transportation. Office of Research, Technology Transfer, and Texas Transportation Institute. 1998. *Development of a methodology to estimate bicycle and pedestrian travel demand.* Research report (Texas Transportation Institute). Texas Transportation Institute, Texas A&M University System. http://books.google.com/books?id=Ql0pAQAAMAAJ.

Venables, W. N., and B. D. Ripley. 2002. *Modern Applied Statistics with S.* Fourth. ISBN 0-387-95457-0. New York: Springer. http://www.stats.ox.ac.uk/pub/MASS4.

Wickham, Hadley. 2009. *ggplot2: elegant graphics for data analysis.* Springer New York. `http://had.co.nz/ggplot2/book`.

———. 2011. "The Split-Apply-Combine Strategy for Data Analysis." *Journal of Statistical Software* 40 (1): 1–29. `http://www.jstatsoft.org/v40/i01/`.