Quality Counts: Developing Accuracy Checks for Bike and Pedestrian Count Data

Nathan McNeil, M.U.R.P., Kristin Tufte, Ph.D.

In the past decade bike and pedestrian count programs have sprung up all over the United States, gathering data to evaluate biking and walking infrastructure. However, these modes have not been studied with the quantitative rigor applied to motor vehicle travel. A research project funded by the National Institute for Transportation and Communities (NITC), led by Nathan McNeil of Portland State University (PSU), offers a method for monitoring the quality of this bike-ped count data.

> "There has been an effort to collect more bike-ped count data in recent years, but it hasn't been consistent in terms of what's being collected and how it's stored. If the data aren't in a uniform format, or aren't stored in a location where they can be easily accessed in bulk, then doing a deep scan of the data would be a challenge," McNeil said.

So what steps need to be taken before data are usable for research, planning or policy?

Researchers and data managers around the country struggle with efficiency versus accuracy. For organizations dealing with large amounts of data where it isn't possible to individually examine each dataset, PSU researchers developed some thresholds of plausibility to flag data that may be erroneous.

This project explores the efficacy of various types of quality checks under different circumstances, including data collected with different types of counters. This project was conducted in partnership with the Oregon Department of Transportation (ODOT). Tammy Lee and Kristin Tufte developed and tested the scripts to check the quality of data from continuous counts.

TYPES OF BIKE-PED COUNTERS

Non-motorized counting methods include:

- **Permanent automatic counters:** infrared beams, pneumatic tubes, inductive loops, magnetometers, piezoelectric strips, radar sensors, optical sensors, and thermal imaging systems; and
- **Short duration counts:** Collected manually by volunteers or by portable automated equipment.

With such a wide variety of methods in place, it is critical to do quality checks to ensure the data are accurate and uniform. The final report, "Biking and Walking Quality Counts: Using "BikePed Portal" Counts to Develop Data Quality Checks" can help agencies clean up data that were collected using any of these methods.

WHAT CAUSES BAD BIKE-PED COUNT DATA?

Researchers in this project focused on identifying data that were inaccurate as a result of equipment failures, such as a counter malfunctioning or going offline, as well as other issues that cause major deviations from expected counts. A variety of environmental factors can also come into play. For example in 2018, the Minnesota Department of Transportation reported a data quality issue from spiders building webs in a bike counter, causing it to report extraneous numbers. Extreme weather, special events and street closures can have a big impact on changes in count data.

THE METHODOLOGY

A key method used by the researchers was to comb through a selection of count data to identify expected

count ranges and patterns, then identify counts on the fringe or tail end of expected ranges or patterns.

Researchers created the framework to be responsive - the expected range can be adjusted and tailored to a location. For example, Portland, Oregon's car-free bridge, Tilikum Crossing, typically sees at least 800 cyclists per day during the off-season (like in mid-November). If the permanent counter on Tilikum Crossing reported zero cyclists on any given day, this would raise a data flag. On a lower-volume street, a report of zero cyclists in a day from that location's bike counter would not necessarily raise any alerts. Flagged data are intended to be reviewed by whoever is most familiar with the dataset, so that a determination can be made about whether the numbers ring true.

> "Our guiding idea was that we wanted the process to be sensitive enough to identify major errors, but also not so sensitive that it overwhelmed people with flags. It was really about, how do you develop this process in a way that is manageable for the person that is uploading or inputting the data? It's striking a practical balance between the user experience and data quality so that the tool actually gets used," McNeil said.

NEXT STEPS

As with all research studies, this exploration into potential quality checks for nonmotorized traffic counts opens the door to more research questions. This includes investigation into potential data quality check differences between pedestrian and bicycle counts. Future work will incorporate weather data, events and street closures into

expected counts and thresholds. In the near future, PSU researchers plan to incorporate the median / interquartile range based adaptive running thresholds to identify site / location specific high and low count thresholds based on day and hourly count.

ABOUT THE AUTHORS

The research team consisted of Nathan McNeil, Kristin Tufte and Tammy Lee, Portland State University; and Krista Nordback, University of North Carolina.

ABOUT THE FUNDERS

This research was funded by the National Institute for Transportation and Communities, with additional support from the Oregon Department of Transportation (ODOT).

THE FULL REPORT and ONLINE RESOURCES

For more details about the study, download the full report **Biking and Walking Quality Counts: Using "BikePed Portal" Counts to Develop Data Quality Checks** at nitc.trec.pdx.edu/research/project/1026

See our Guide to Bicycle and Pedestrian Count Programs at: **trec.pdx.edu/resources-and-tools/#5228/Guide_to_Bicycle_Pedestrian_Count_Programs**

Photo by Transportation Research and Education Center at Portland State University