Leveraging GTFS data to assess transit supply

James Reynolds[a,1,*], Yanda Qu[a,3], Graham Currie[a,1]

[a]*Institute of Transport Studies, Department of Civil Engineering Engineering, Monash University, Clayton Campus, Melbourne, 3800, Victoria, Australia*

Abstract

This is the abstract.
    It consists of two paragraphs.

Keywords: keyword1, keyword2

1. Introduction

Berenson (2016) highlights that the quote "if you can't measure it, you can't manage it" is often miss-attributed to Deming (1993), and that he was actually trying to make the opposite point in *If you can't measure performance, can you improve it?*. Regardless, service level indicators are an important part of researching, managing and seeking to improve transit operations (Fielding, 1987; Ryus et al., 2003). A wide range of indicators already, including, for example: those in the Transit Capacity and Quality of Service Manual (TCQSM)(Kittleson & Associates et al., 2013), the Transit Score metric (Walk Score, 2023) and many more.

Practitioners, researchers and advocates seeking to use such metrics may face two inter-related challenges: (1) calculating the metrics themselves for a specific location and service pattern; and (2) explaining the metrics, their meaning and importance to those who might not be specialists in transit, such as to politicians or the general public. For example, the TCQSM metrics appear difficult to calculate in practice without access to specialist software and data. They appear relatively easy to explain given they use an A to F scoring system and there is an entire guidebook about them. However, although this may be offset by the multitude of different indicators. In contrast, Transit Scores can be obtained simply by typing an address into a website, but cannot be calculated independently as the methodolgy / algorithm is not publically available.

Previous research by Currie and Senbergs (2007) developed a transit Supply Index (SI) metric that is both relatively easy to calculate and explain. It is obtained by calculating the number of transit arrivals at stops within an area of interest, with an adjustment made to account for the typical walk-access catchment for each stop. Hence, higher SI scores indicate areas with higher frequency and/or better coverage.

Unfortunately, the SI does not appear to have been widely used, perhaps in part because at the time it was first published timetable data was not publicly available in a standardized and machine-readable format. The scores reported in Currie and Senbergs (2007) were calculated directly from a database of services provided by the transit authority in Melbourne, Australia. Since then, however, the General Transit Feed

---

*Corresponding author
    Email addresses: `james.reynolds@monash.edu` (James Reynolds), `yanda.qu@monash.edu` (Yanda Qu), `graham.currie@monash.edu` (Graham Currie)
    [1]Research Fellow
    [2]PhD Strudent
    [3]Professor

Specification (GTFS) has been developed as a way to publish timetable data, along with many processing tools. More than 10,000 agencies are now providing GTFS feeds[4] (MobilityData, undated),

A gap, however, is that a tool to calculate SI scores directly from a GTFS dataset has not yet been available. This provides the motivation in this paper, in which a new R package (gtfssupplyindex) specifically developed to calculate SI scores is presented.

References

Robert A. Berenson. If you can't measure performance, can you improve it? *JAMA*, 315(7):645–646, 02 2016.

Graham Currie and Zed Senbergs. Identifying spatial gaps in public transport provision for socially disadvantaged Australians: the Melbourne needs-gap study. In *Australasian Transport Research Forum*. Australasian Transport Research Forum, 2007.

W Edwards Deming. The new economics for industry, government, education, 1993.

Gordon J Fielding. *Managing public transit strategically: a comprehensive approach to strengthening service and monitoring performance.* Jossey-Bass public administration series. Jossey-Bass Publishers, San Francisco, 1st ed. edition, 1987. ISBN 1555420680.

Kittleson & Associates, Parsons Brinckerhoff, KFH Group, Texas A&M Transportation Institute, and ARUP. *Transit Capacity and Quality of Service Manual, Third Edition.* Transportation Research Board, Washington DC, third edition, tcrp report 165 edition, 2013. URL http://www.trb.org/Main/Blurbs/169437.aspx.

MobilityData. *General Transit Feed Specification (GTFS)*, undated. URL https://gtfs.org/.

Paul Ryus, M Connor, S Corbett, A Rodenstein, L Wargelin, L Ferreira, Y Nakanishi, and K Blume. Tcrp report 88: a guidebook for developing a transit performance-measurement system. Technical report, 2003.

Walk Score. Transit score methodology. 2023. URL https://www.walkscore.com/transit-score-methodology.shtml.

---

[4]There are two forms: GTFS-static consisting of the timetable data (the scheduled services); and GTFS-realtime, which includes vehicle arrivals and departure times based on real-world position data. This paper and project uses only the GTFS-static (timetable) format.