

Appendix D:

Methodology

This section outlines the methodology used for the complex data analysis in bus and train timetable integration and households' distance from bus and public transport services.

Benchmarks for bus and train timetable integration

Data sources

We use daily timetable data for metropolitan buses and trains. Timetable data is provided by the department and used by VAGO to evaluate the transfer window between connecting services arriving and departing.

We use a list of priority connections provided by the department to select high priority bus route to train station connections. Bus to bus connections, bus to school bell connections, and bus to V/Line connections are out of scope for this analysis. This data includes the parent route and direction of buses, but not the connecting bus stop ID or name. We use bus stop name data from daily timetables to identify bus stops named after connecting stations.

Sampling method

We select a small representative sample of 12 bus–train connections to evaluate for a 2-year period from July 2023 to June 2025. We evaluate a larger sample of more than 200 bus to train connections over 4 weeks in June 2025.

The larger sample includes only high priority connections and we have been able to confirm a bus stop connecting to the train station using bus stop name. We further filter the sample of connections to bus routes and stations which connect in both the morning and the evening so we have the same number of connection points for each test.

The small sample of 12 is extracted from the larger connections sample using random sampling with 6 strata. It includes 2 parent routes from each of the main 4 bus operators (CDC, Dysons, Kinetic and Ventura) and 2 parent routes from each of the 6 metro regions (Eastern, Inner Metro, Inner South East, Southern, Northern and Western).

Figure D1: The sample of 12 representative bus–train connections

Region	Bus operator	Parent route	Train stations
Eastern	Kinetic	270	Box Hill
Eastern	Kinetic	370	Mitcham
Inner Metro	Dysons	504	Clifton Hill
Inner Metro	Dysons	505	Royal Park
Inner South East	CDC	612	Glen Iris
Inner South East	Kinetic	603	Elsternwick
Northern	Dysons	570	Thomastown
Northern	Ventura	551	Heidelberg
Southern	Ventura	699	Tecoma
Southern	Ventura	825	Moorabbin
Western	CDC	461	Watergardens
Western	CDC	412	Laverton

Source: VAGO.

Analysis method

A connection between buses and trains is within the transfer window if:

in the morning ...	in the evening ...
an inbound train departs 5 or more minutes later than a bus arrives but not more than 10 minutes later.	an outbound train arrives 5 or more minutes earlier than a bus departs but not more than 10 minutes earlier.

Note: For Laverton station, the times are 7 and 12 minutes instead of 5 and 10 minutes.

We analyse connections for both morning and evening and for both weekdays and weekends.

Morning connections are bus services of the identified parent route arriving at a train station at or after 5am but before 3pm and connecting to any train departing inbound toward the city.

Afternoon connections are outbound train arrivals connecting to bus services of the identified parent route that depart at or after 3pm until 5am the next morning.

Both weekdays and weekends have the same definition for morning and afternoon services. Night bus services will also be categorised as morning or afternoon according to this definition.

We do not have data on train outages. We instead infer blocks of time when trains services are not running, to prevent planned train outages unfairly reducing the proportion of bus services that connect to trains.

We exclude 6-hour blocks of time where there are zero train services, or the number of train services drops significantly below usual service numbers, for a train station.

For each bus service arriving at or departing from a train station, we identify whether any train meets the conditions for a connection. Our results return the proportion of connections made per the number of bus services.

Households' distance from bus and public transport services

Data sources

We use publicly available information from DataVic for addresses, roads and trails data, and public transport stops to generate distances between residences and nearby public transport services.

We filter roads to exclude non-walkable roads, for example freeways, ferry routes and planned roads. We do not consider other road accessibility features such as road-crossing points or terrain.

We filter address data to reduce the number of locations that are not accessible by road, non-residential or still being developed in June 2025. We also exclude address data with features that indicate lower quality source data or duplicated addresses. Public transport stops included are metro bus, metro tram and metro train.

We use the department's daily timetable data to determine bus services' quality at bus stops during June 2025.

A bus stop is considered to have a high quality service if there is at least one bus route meeting the high quality service conditions at that specific location for a majority of the month. Only metropolitan bus routes are considered.

Special bus services and FlexiRide services in the metropolitan area are out of scope.

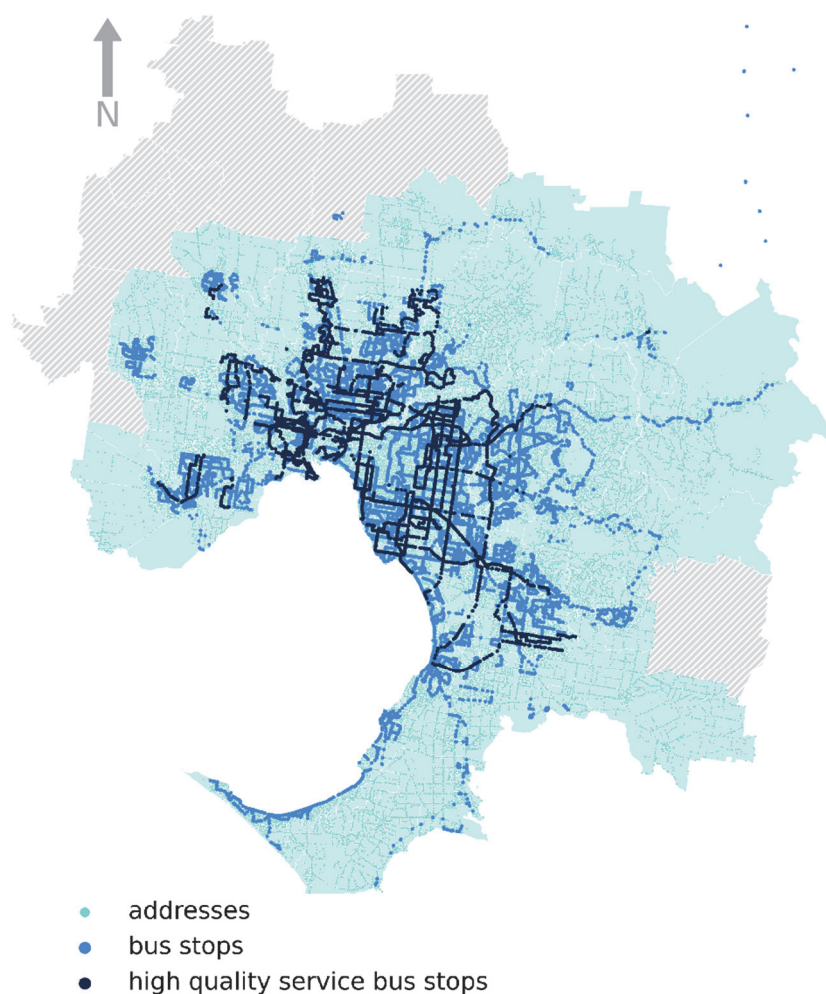
We found 31 per cent of bus stops had at least one out of 366 for metropolitan bus service routes that met these conditions in June 2025. Tram and train stops are all considered have a high quality service for this analysis.

We use projected population and number of households from Victoria in Future data to aggregate data from Statistical Areas Level 2 (SA2) into local government areas and the greater Melbourne region.

Analysis method

Spatial data for addresses, public transport stops and roads is filtered to SA2 regions within greater Melbourne. For each area, we also include roads and public transport stops within a buffer zone of one kilometre outside the area boundary. Seven SA2 regions with no public metropolitan public transport, and one region (Wallan) served more by regional bus services than metropolitan bus services, have been excluded from our analysis. The extent of the addresses within the region and the public transport network is shown in Figure D2.

Figure D2: Addresses within the greater Melbourne area covered by metropolitan public transport services



Source: VAGO.

Within each region, we randomly sample 20 per cent of addresses. For each of the selected addresses, we identify the 10 closest stops in a straight line for:

- nearby metropolitan public transport stops
- nearby metropolitan public transport stops with a high quality service
- nearby metropolitan bus stops
- nearby metropolitan bus stops with a high quality service.

The sample of 20 per cent of addresses is a large enough number to be representative. The selection of the closest 10 stops in a straight line across different criteria is enough to provide confidence that the closest stops along the road network are in the selected address-stop pairs.

For each unique address-stop pair, we find the distance along the road network Dijkstra's algorithm.

Our method first requires a calculation of the distance from an address to the nearest midpoint on the nearest road and preparation of the network so that the nearest midpoint is converted into a vertex for accurate routing.

The routing algorithm for each address-stop pair is solved independently of other address-stop pairs.

We use parallel processing, with batches of 1,000 addresses, for each SA2.

We sum the distance from an address to the nearest road and the distance along the road to the nearby stop to get the distance from an address to the nearby stop. For each address, we create flags to indicate whether there is at least one nearby stop that is:

- within 400 metres and provides access to a metropolitan bus service
- within 400 metres and provides access to a metropolitan bus, train or tram service
- within 800 metres and provides access to a high quality metropolitan bus service
- within 800 metres and provides access to a high quality metropolitan bus service, train service or tram service.

Our results show the percentage of addresses for each SA2 that have accessible public transport. We use projected population in 2025 to weight SA2 data when aggregating to larger areas.
