Improving the Bikeability of Our Cities

(Cycling Infrastructure Scenario Builder)

Final Project Report

February 2024

Professor Chris Pettit

Associate Professor Brian Lee

Dr Hao Wu

Dr Lee Roberts

City Futures Research Centre
School of Built Environment | University of New South Wales

UNSW City Futures Research Centre

UNSW SYDNEY

Executive Summary

With increased cycling participation in Australian cities and the growing urgency to deliver more cycling infrastructure to accommodate cycling, better decision support tools are needed to inform where to build these new cycling infrastructure. With funding from the Australian Research Council (ARC) and in collaboration with NSW Office of Sport, Penrith City Council, Transport for NSW, Wollongong 2022 Limited, and Wollongong City Council, the *Cycling Infrastructure Scenario Builder* has been developed by the City Futures Research Centre at the University of New South Wales to provide decision support for cycling infrastructure planning.

The *Cycling Infrastructure Scenario Builder* is a web-based interactive decision support tool that is designed to facilitate the planning of cycling infrastructure. The tool enables users to test various cycling infrastructure scenarios on digital maps and estimate cycling uptake resulting from hypothetical new cycling infrastructure. The tool incorporates a wide range of cycling-related spatial maps and data layers to provide contextual information and visual guidance, which enhances the interactive experience while planning for cycling infrastructure.

This tool adopts a data-oriented approach in modelling and estimating the would-be effect of hypothetical cycling infrastructure. Estimates of cycling uptake is produced by a discrete choice model, which was calibrated using data from a recent (2022-23) cycling survey conducted as part of this project in Greater Sydney. The discrete choice model predicts how individual level cycling participation would change in response to changes in cycling infrastructure and improvement in access to points-of-interest (POIs). This model is combined with a synthetic population model of the Greater Sydney in order to predict cycling participation.

This tool has a wide range of applications. Existing data layers allow the visualisation of existing cycling infrastructure, existing cycling behaviours and natural and built environment characteristics that influence cycling. Existing crash data can be used to identify problem areas and help identify safety hazards. The scenario tool enables the planning of strategic cycling networks, testing and evaluation of proposed facilities from local or state government plans, identification of health and economic benefits of cycling infrastructure provision, and prioritisation of cycling infrastructure projects. A streamlined batch-processing mode allows rapid testing of a large number of infrastructure scenarios, saving time and resources for users of this tool. The tool synthesises cycling related data from various sources within a single platform, and can also be useful for training and educational purposes. The tool is designed to be extensible to enable new data layers, new transport modes such as walking, and new jurisdictions outside the current range. This document provides a detailed account of the functionalities of the tool, and the under-the-hood modelling methodology.

Table of Contents

1. Dashboard Design and Functions

1.1 Background

In Greater Sydney, Australia, a large proportion of the population are either current cyclists or are interested in cycling (Munro, 2023). A significant amount of new cycling infrastructure has been built to accommodate and encourage the increase in cycling. More cycling infrastructure is being planned. There are challenges in the process of planning for cycling infrastructure. In particular, it has been both a challenge and an opportunity for practitioners at local and state levels to understand to make the best use of investment in cycling infrastructure and encourage more people to ride bikes. Comparing the effectiveness of different cycling infrastructure scenarios and predicting the effect of new infrastructure in increasing cycling uptake would benefit from better data-oriented decision support tools.

While more and better cycling infrastructure is generally associated with higher cycling participation, the extent of cycling uplift depends on a wide range of other factors. The adjacent population density determines the number of people that can potentially use the cycling infrastructure; social demographics and local culture (sentiment) also affect the likelihood that an individual will choose to ride a bike. The likelihood to cycle is further affected by the built environment, including the presence or absence of urban amenities and existing levels of cycling infrastructure. Given the wide range of factors, the same cycling infrastructure investment would be more effective in generating new cyclists in some areas compared to others. This tool has been developed to inform and support decision making around 'where to build' new cycling infrastructure.

Strategic planning considerations, such as the need to integrate new cycling paths into the existing cycling network, connect known cycling hot spots, fill gaps in the network and avoid hilly terrain and dangerous areas, are also important in effective provision of cycling infrastructure. Supplying contextual information about the local built environment and providing practitioners with the ability to plan, visualise, and estimate the effects of new cycling routes would help with better decision making.

The *Cycling Infrastructure Scenario Builder* is intended to deliver an integrated data-oriented decision support tool that visualises cycling related information on GIS-based maps, allows users to rapidly test different cycling infrastructure scenarios, and produces estimates for the resulting changes to cycling participation. The tool accounts for complex factors and relationships that affect the effectiveness of new cycling infrastructure. The tool currently covers existing cycling infrastructure and the adult population in Greater Sydney, but the data and modelling method of this tool are scalable, and could be extended to include future and proposed cycling infrastructure, and to extend the tool to other cities and areas.

The development of this Cycling Infrastructure Scenario Builder was funded by the Australian Research Council (ARC), and in collaboration with NSW Office of Sport, Penrith City Council, Transport for NSW, Wollongong 2022 Limited, and Wollongong City Council. In order to ensure that the final product meets the needs of potential users, the project team sought feedback from stakeholders throughout the development process, and iteratively improved the tool.

1.2 Functionalities of the tool

The primary purpose of this tool is to assist in the planning of cycling infrastructure and to account for complex factors and relations that affect cycling uptake. Functionalities of the tool are designed around this purpose. Specifically, the tool integrates three major functions, namely

- **Explore current conditions** Explore current conditions and visualise cycling related information on GIS map layers
- **Create changes via interactive design and visualisation** Allow users to interactively create changes, and visualise new cycling infrastructure on GIS map layers
- **Estimate uptake in cycling participation** Estimate uptake in cycling participation from user-envisioned cycling infrastructure scenarios

Each of the three functions are discussed in detail in this section. These functions serve an important role in guiding users through the creation of new cycling infrastructure scenarios, and in providing feedback for the effectiveness of new cycling infrastructure. The flow chart in Figure 1 shows the function of different components of the tool, and the interactive process of using the tool for cycling infrastructure planning.

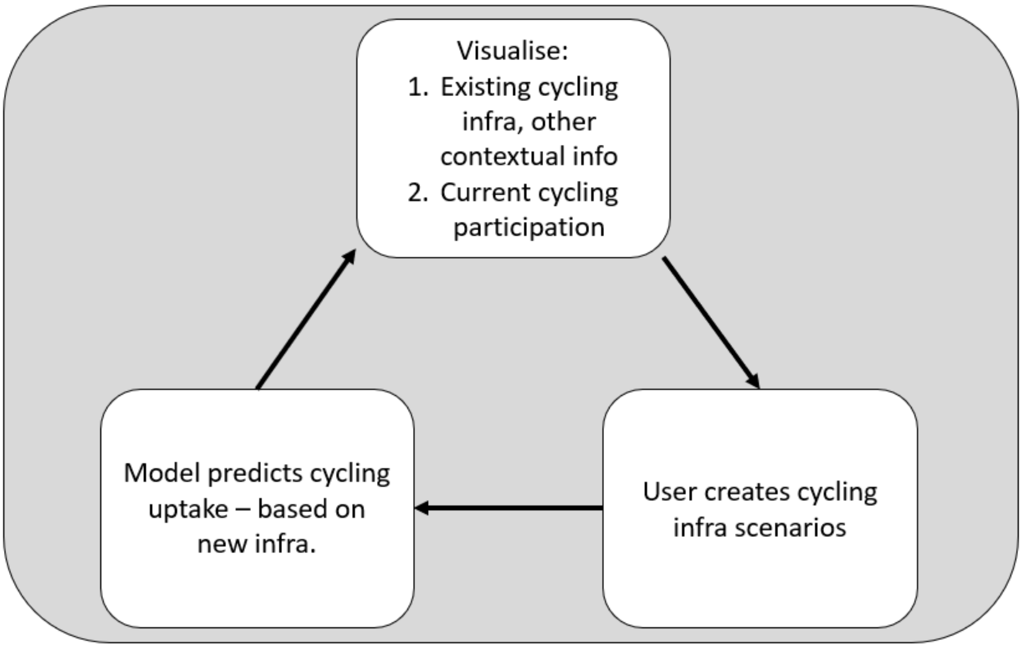

Figure 1. Flowchart showing the function of different components of the tool

The tool is hosted on a webpage with interactive GIS map displays. The map allows the visualisation of existing cycling infrastructure and cycling related information, which provides background information for users to create cycling infrastructure scenarios. Users are then able to create new infrastructure scenarios by selecting road segments to add or change different types of cycling infrastructure. User-added cycling infrastructure is then evaluated by the model to evaluate the impact.

This tool is designed to be easily accessible through a web user interface to provide cross-platform compatibility. Users can access the tool through any device with an internet browser and a connection to the internet.

Explore current conditions

The “Explore current conditions” function is intended to familiarise users with the existing context around a proposed cycling corridor, and to improve situational awareness. In order to achieve higher cycling uptake, the planning of new cycling infrastructure needs to consider many factors that are contextual to the local built environment. Namely, the placement of new cycling infrastructure needs to be easily reachable by a sufficient number of people, especially those who view cycling positively and have a high likelihood to take up cycling if appropriate infrastructure was provided. The new cycling infrastructure also needs to provide connections to common destinations (referred to as points of interest, or POIs, in this tool). Other considerations for new cycling infrastructure include the relationship with existing cycling infrastructure, and the need to avoid unfavourable terrain and busy roads. In order to plan for cycling infrastructure, practitioners often need to collect data and synthesise information, which can be both time consuming and resource intensive.

The ”Explore current conditions” function visualises existing cycling infrastructure on GIS maps, along with the location of cycling related amenities and dis-amenities to provide important contextual information and guidance in the placement of new infrastructure. The tool integrates a wide range of data in order to provide users with useful information and visual guidance. Information such as existing road network and cycling infrastructure, spatial distribution of jobs, population and urban amenities are shown in GIS maps as different layers using lines, shapes and heatmaps, which improves situational awareness as users plan new cycling routes.

Users are able to toggle each layer on and off. Some of these features, such as steep slopes, are disfavoured by cyclists, and therefore should be avoided when possible when planning for new cycling infrastructure. Other features such as favourable cycling sentiment, high population density, abundance of POIs and job locations create ideal conditions for high cycling participation, and additional cycling infrastructure at these locations will tend to contribute more to cycling uptake. Gaps in existing cycling infrastructure can be identified, as well as places with conducive conditions for cycling, but where cycling infrastructure is lacking. Based on contextual information provided by this tool, users will be able to apply their own criteria in locating and formulating new designs to improve cycling infrastructure. Table 1 details each of the cycling related GIS map layers, and example images of each layer are shown in Figure 2 through Figure 15.

Data Layer	**Definition**	**Source**
Satellite imagery	Base satellite imagery layer for contextual information	Google
Cycling sentiment	Percentage of local government area (LGA) residents with favourable attitudes cycling; based on cycling survey data	CFRC
Cycling participation (by cycling purposes)	Percentage of LGA residents cycling for different purposes; based on cycling survey data	CFRC
Severe cycling crashes	Location of traffic accidents involving cyclists (from NSW police reports)	TfNSW

Base road network	Road network based on OpenStreetMap (OSM). This road network layer is used for interactive design of new cycling infrastructure	OSM (Crowd-sourced)
Existing cycling infrastructure	Based on TfNSW bicycle network data	TfNSW
Cycleway network connectivity	This layer shows isolated segments of the cycling infrastructure	CFRC
Edge to node ratio	The ratio between the number of edges (roads) to the number of nodes(intersections); a higher ratio suggests a better connected network	CFRC
Slope of road network	Gradient of each road segment.	Digital Elevation Data - CFRC
Population density	Density of residential population	ABS
Points of interest (POIs)	Parks, schools, service businesses, shopping businesses, train stations	OSM
POIs within catchment	The number of points of interest within user-defined catchment distance	OSM -CFRC
Jobs within catchment	Number of jobs within user defined catchment distance	ABS -CFRC
Mesh Blocks with school connectivity	Mesh Blocks that are connected to local catchment public high school(s) via safe bicycle infrastructure.	CFRC

Table 1. Integrated data layers for cycling related contextual information

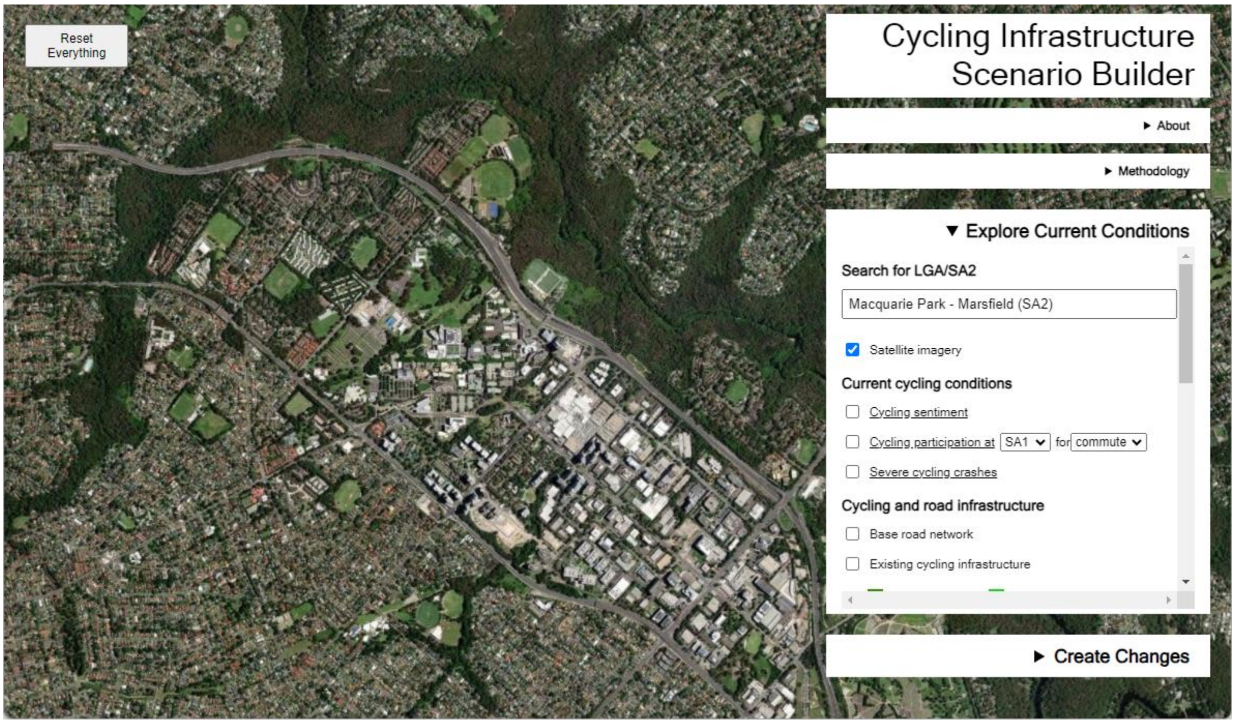

Figure 2. Satellite imagery

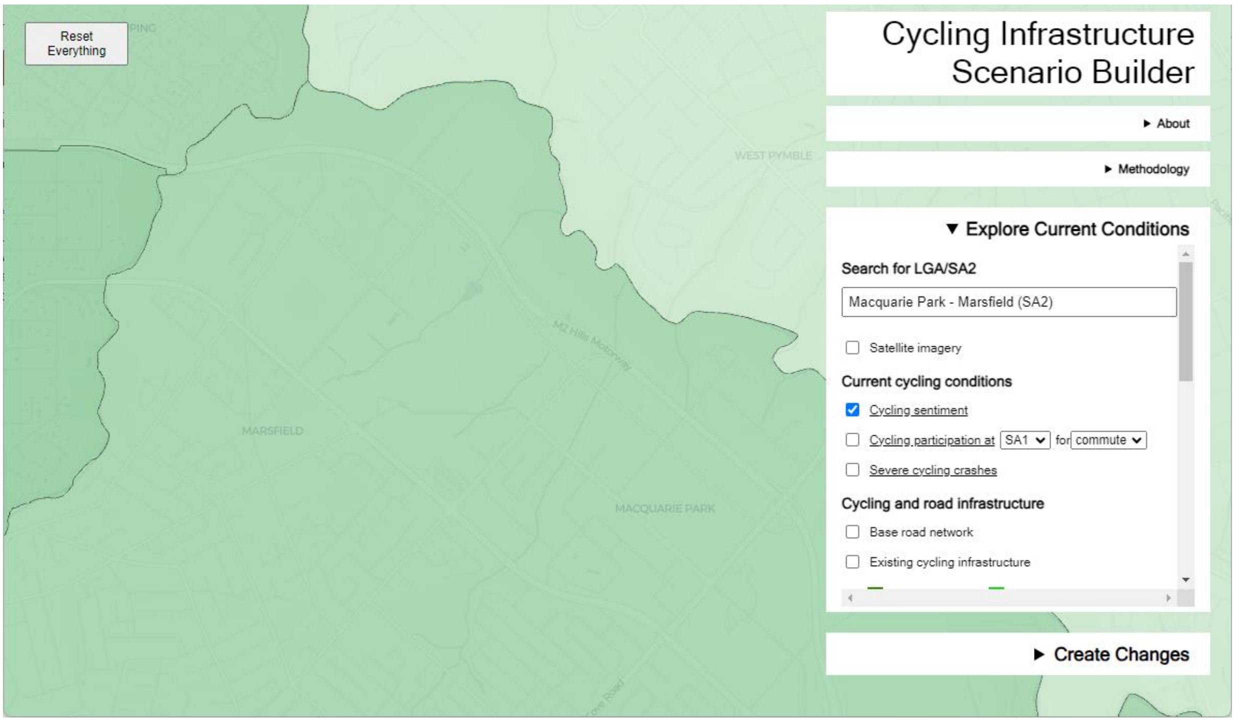

Figure 3. Cycling sentiment

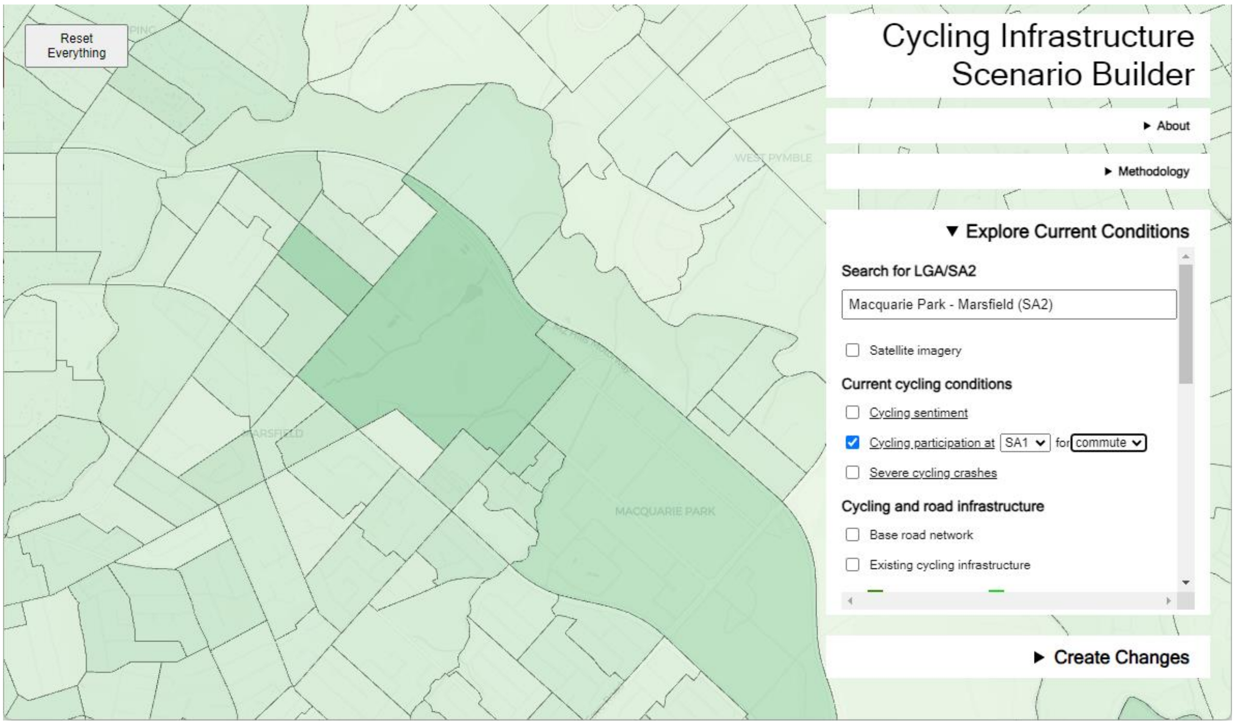

Figure 4.Cycling participation (for commute purposes at SA1 level)

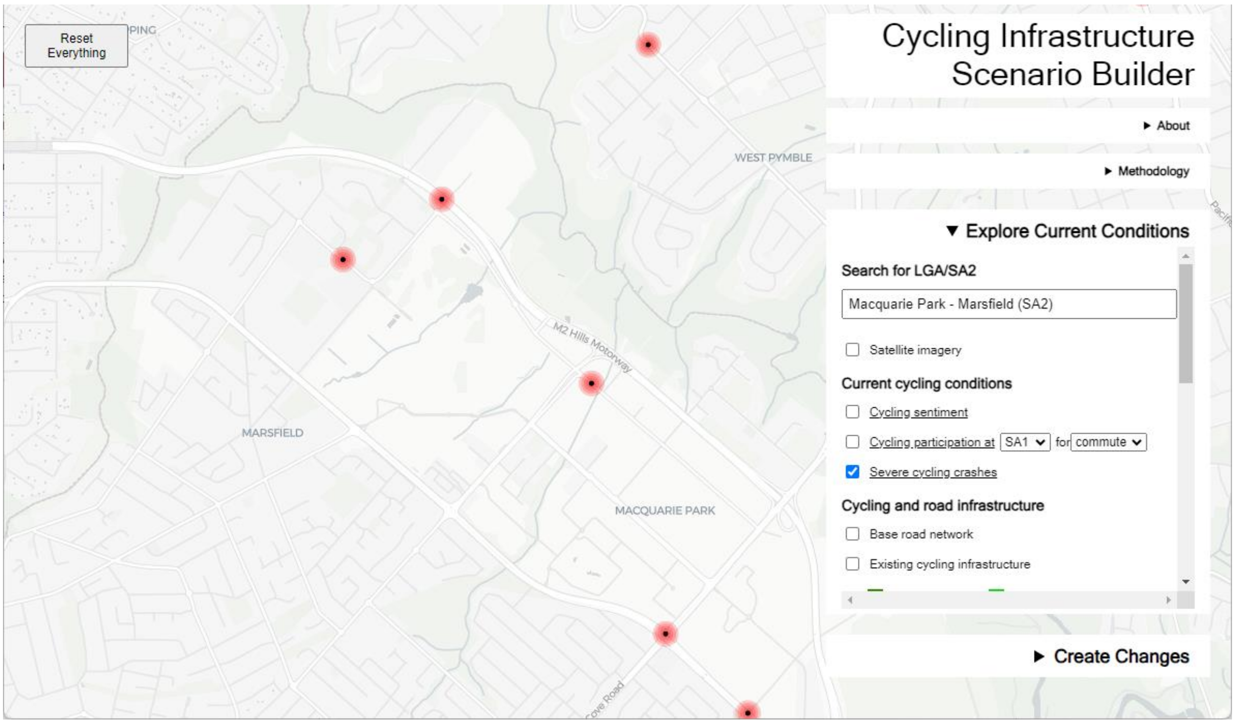

Figure 5. Severe cycling crashes

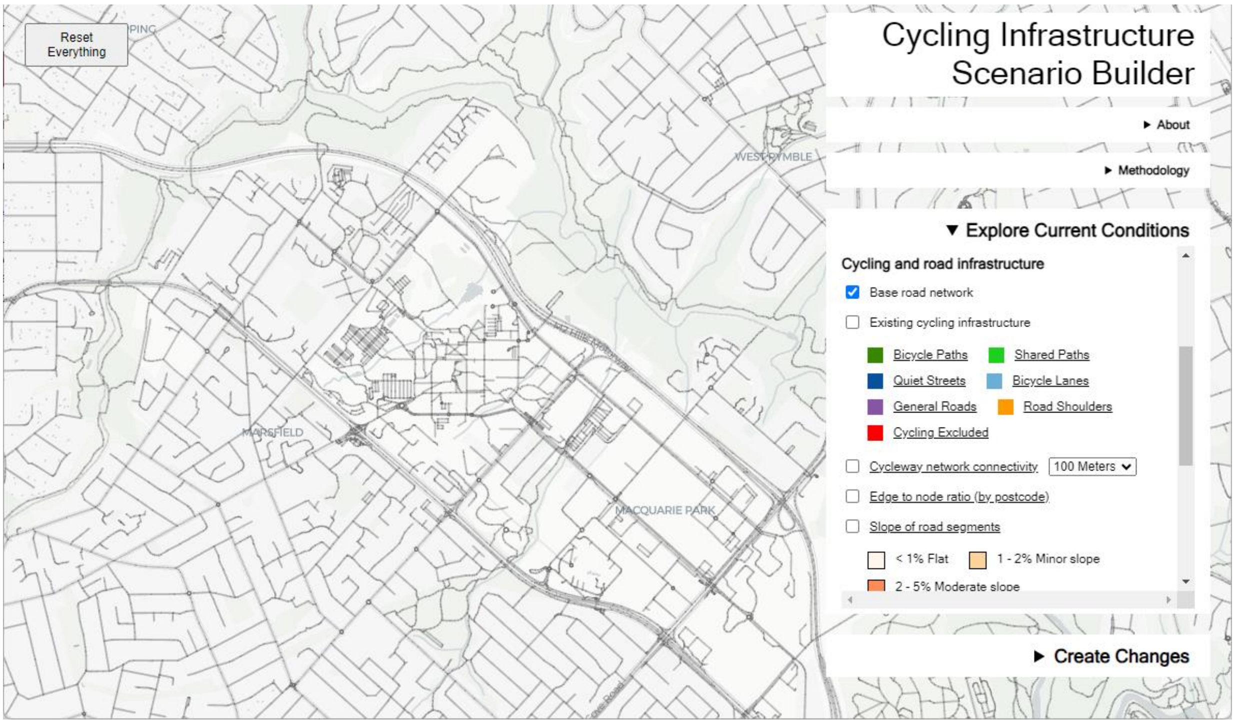

Figure 6. Base road network

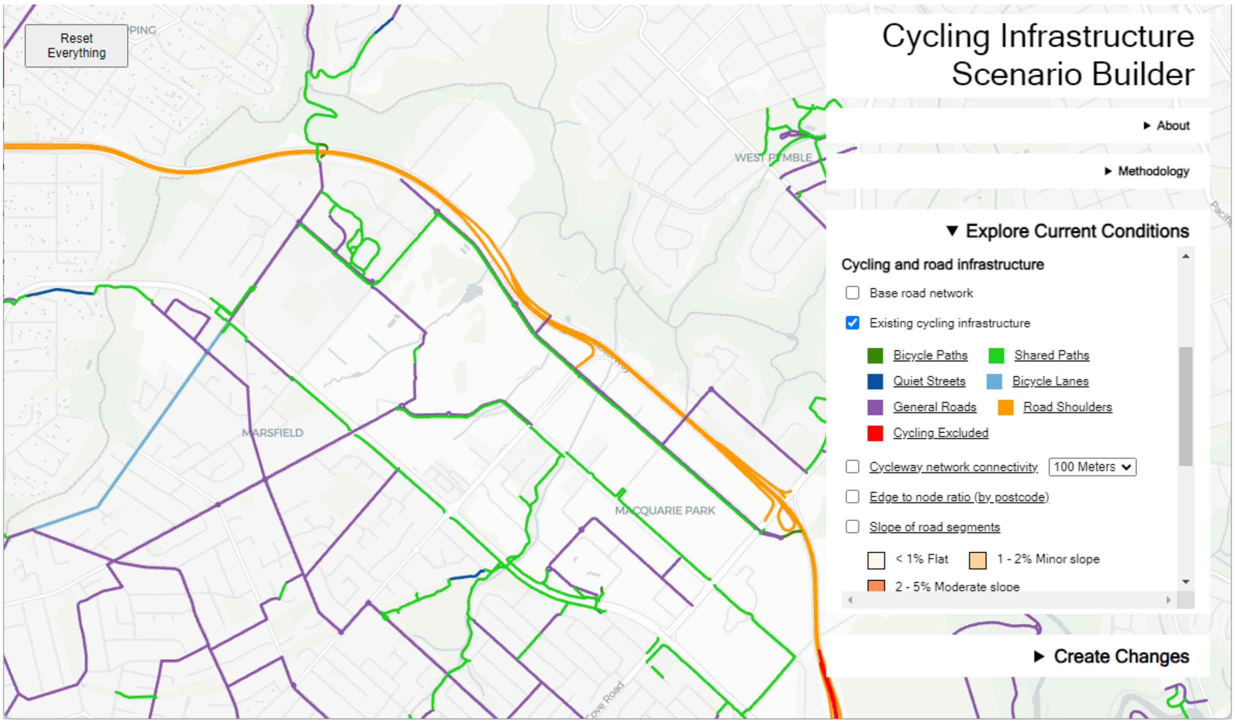

Figure 7. Existing cycling infrastructure

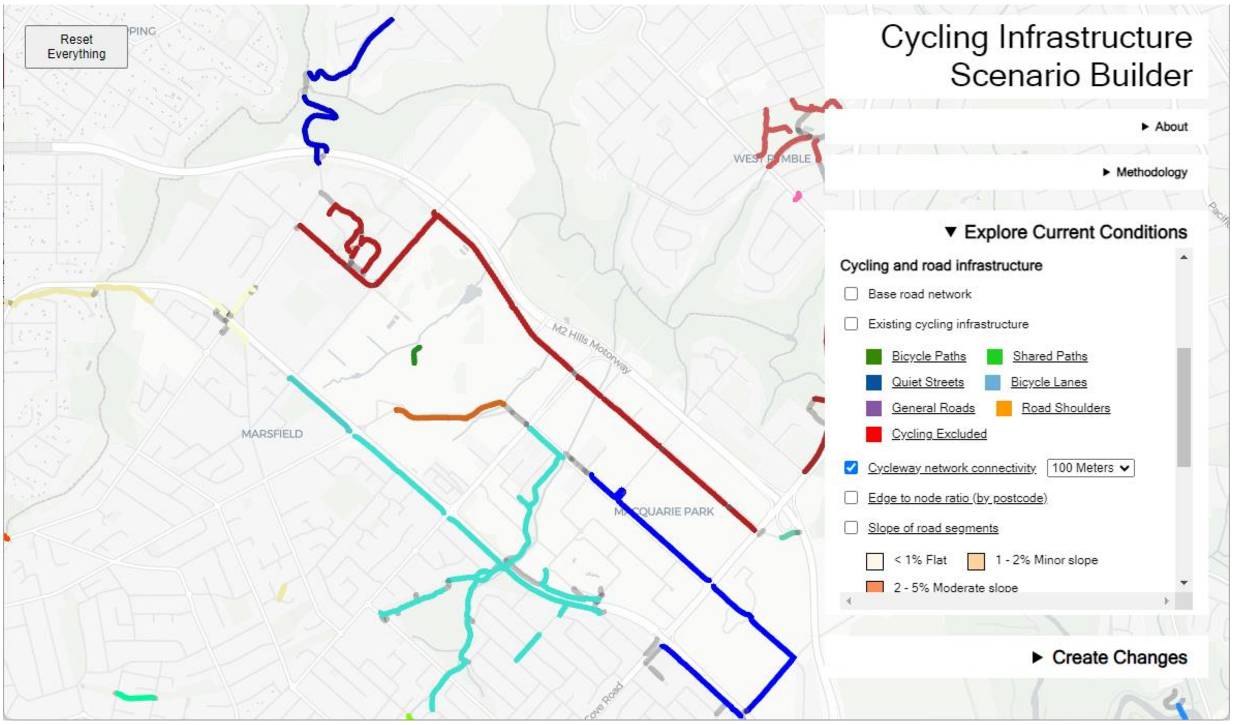

Figure 8. Cycleway network connectivity

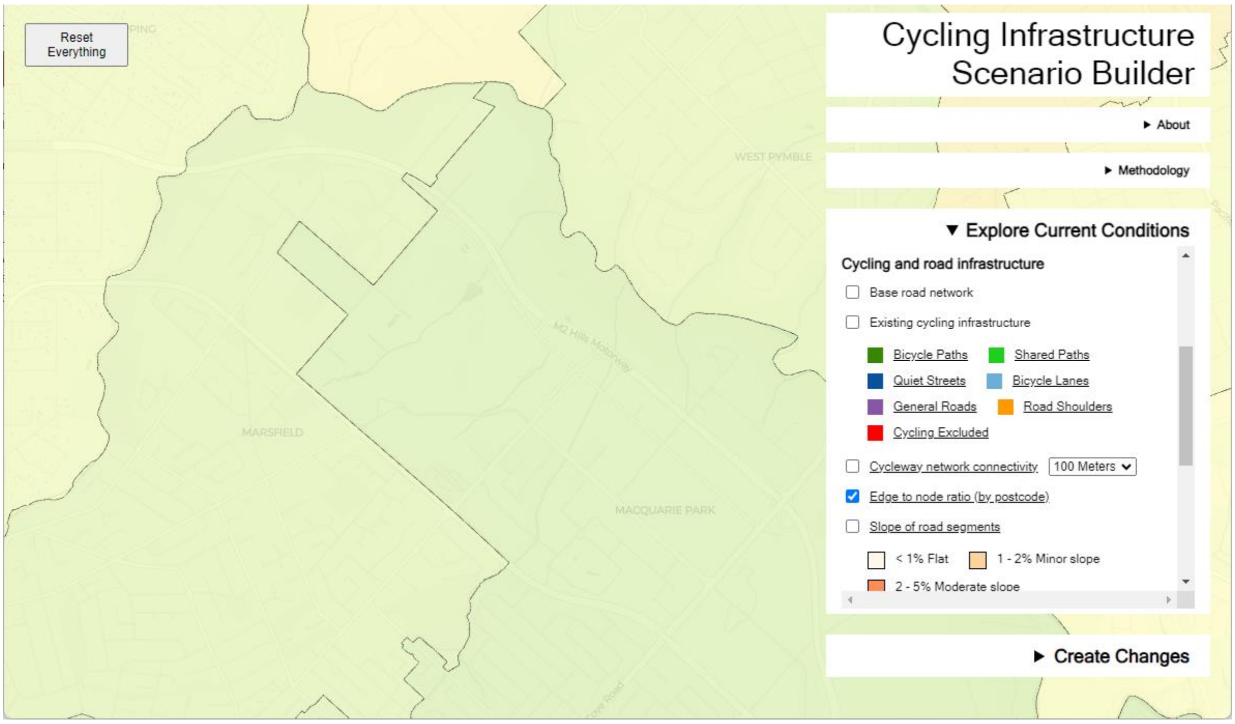

Figure 9. Edge to node ratio

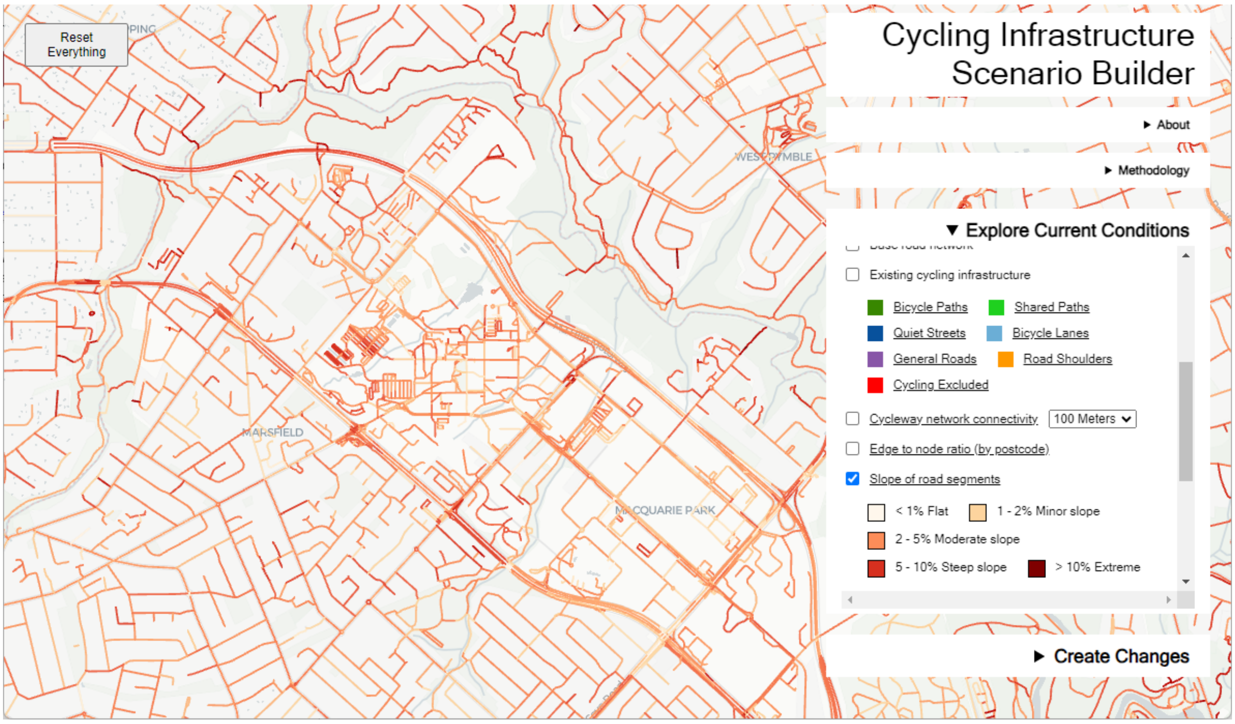

Figure 10. Slope of road network

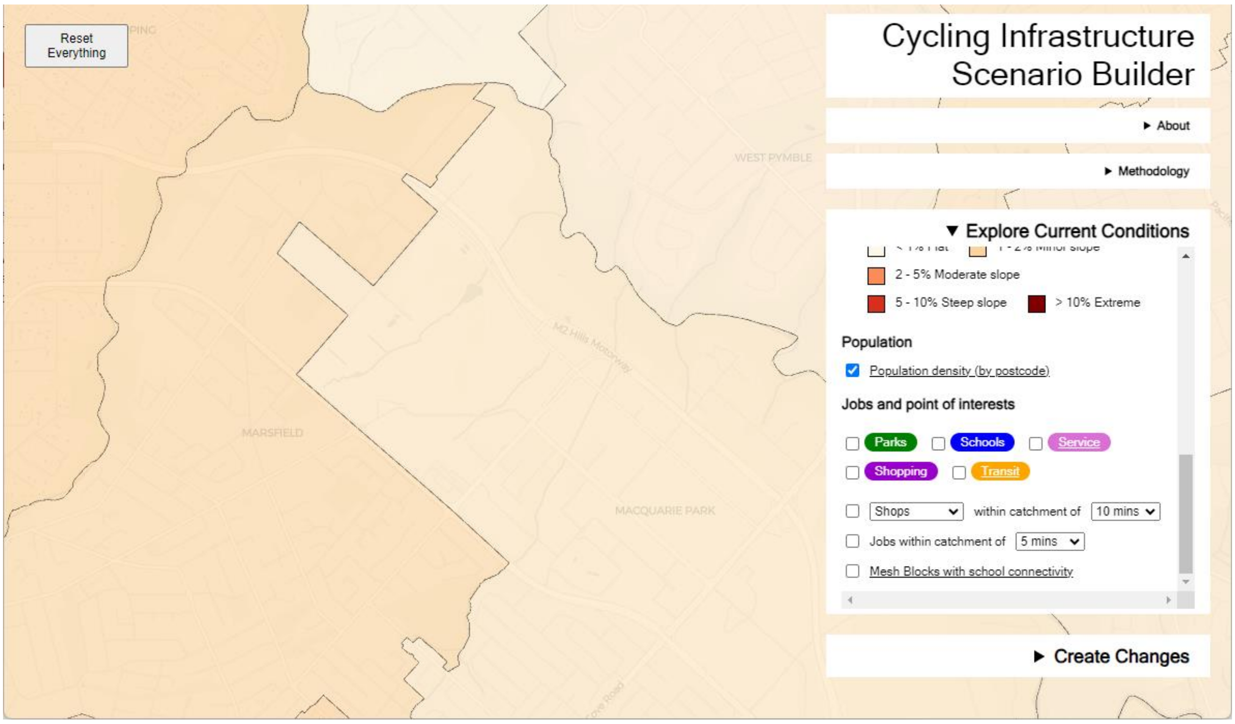

Figure 11. Population density

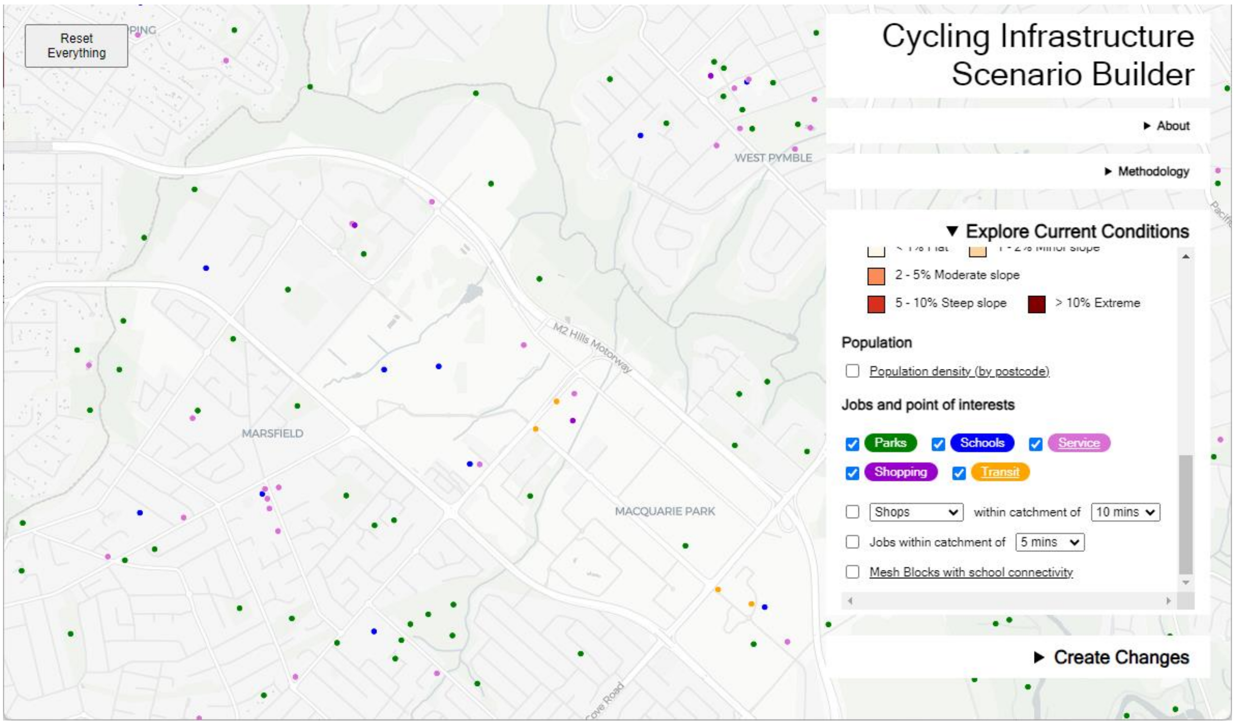

Figure 12. Points of interest (POIs)

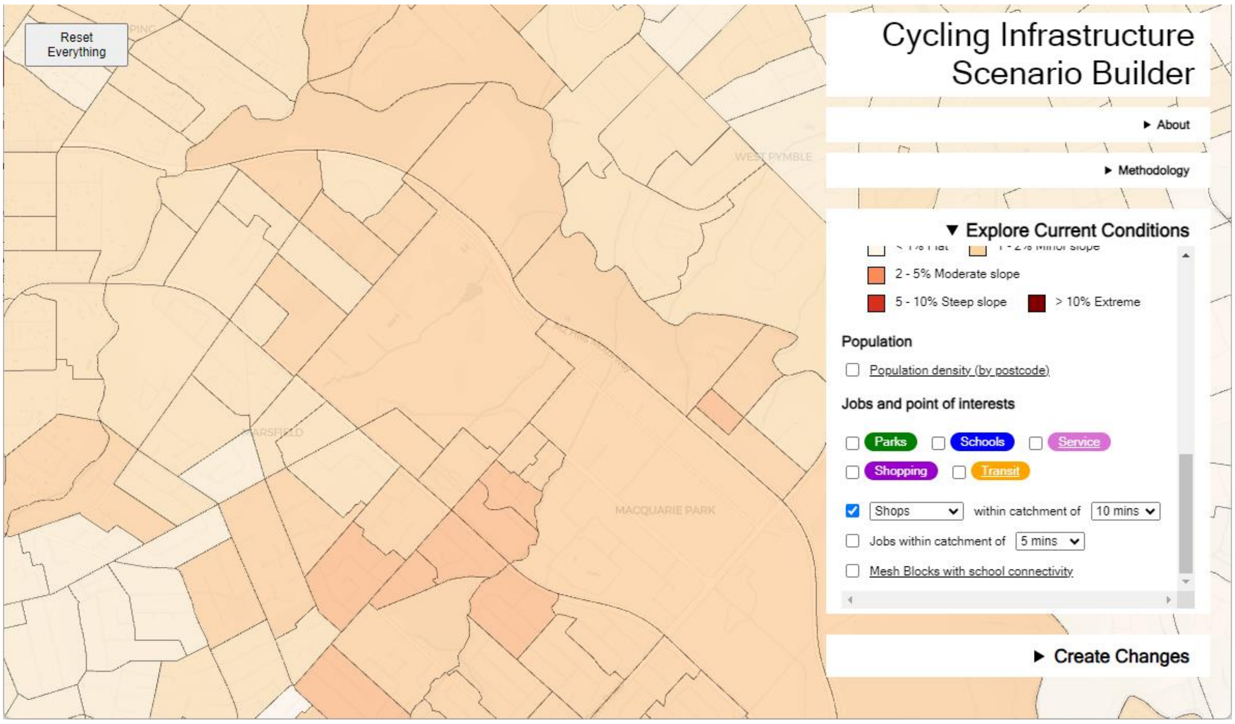

Figure 13. POIs within catchment (e.g., shops within 10 minutes by bicycle from each SA1)

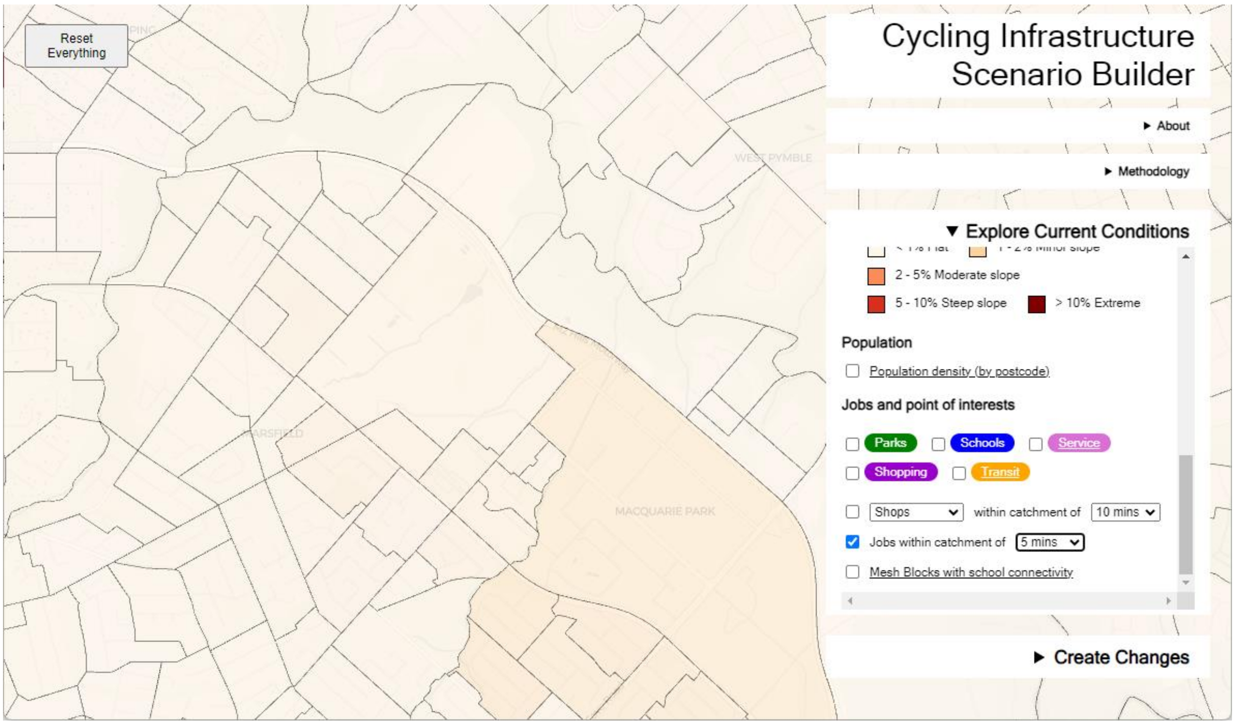

Figure 14.Jobs within catchment (e.g., 5 minutes by bicycle from each SA1)

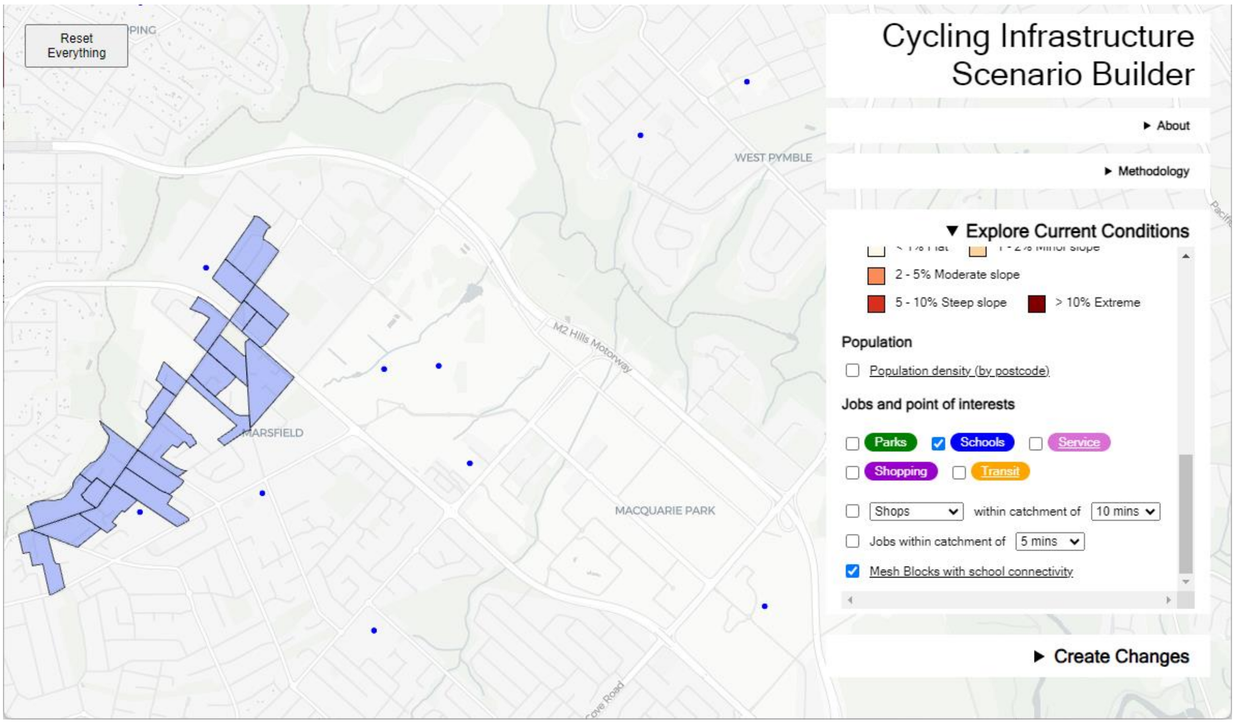

Figure 15. Mesh blocks with cycleway connectivity to catchment public school

Create changes, interactive design and visualisation

After exploring the current conditions and cycling related information, a starting point of using this tool for cycling infrastructure planning is to create and visualise designs for new cycling infrastructure on GIS map layers. Through the interactive web user interface, users are able to select road segments and add or change cycling infrastructure types on selected segments. Roads with different types of cycling infrastructure are visualised in different colours. Figure 16 shows an example of selecting and visualising a new segment of cycling infrastructure (in red) through the interactive user interface.

The ability to interactively create and visualise what the new route looks like, what origins and destinations are better connected by the new route, and how the new route enhances existing cycling networks provides a wide range of possible uses for this tool.

The tool also includes a batch processing mode to allow the rapid testing of a large number cycling infrastructure scenarios. In the batch processing mode, instead of manually selecting and defining the type of cycling infrastructure on each segment, information about new cycling infrastructure can be imported via a file upload interface. Results from model prediction are exported in JSON format. In order to use the batch processing mode, the location and type of new cycling infrastructure need to be prepared in GeoJSON format. Existing infrastructure data in (legacy) ESRI Shapefile format can be easily converted for batch processing, which provides good compatibility with existing data that might be available to users. A wide range of GIS software (QGIS, ArcGIS) are available for either converting ESRI shapefiles, or creating GeoJSON data for batch processing.

The ”create changes” function is based on the OpenStreetMap (OSM), which allows users to add or modify cycling infrastructure on existing road segments. Under certain circumstances users may wish to build new cycling infrastructure scenarios at places where no existing road exist. This can be achieved through the batch processing mode, since the batch processing mode does not require new cycling infrastructure to be on existing road segments.

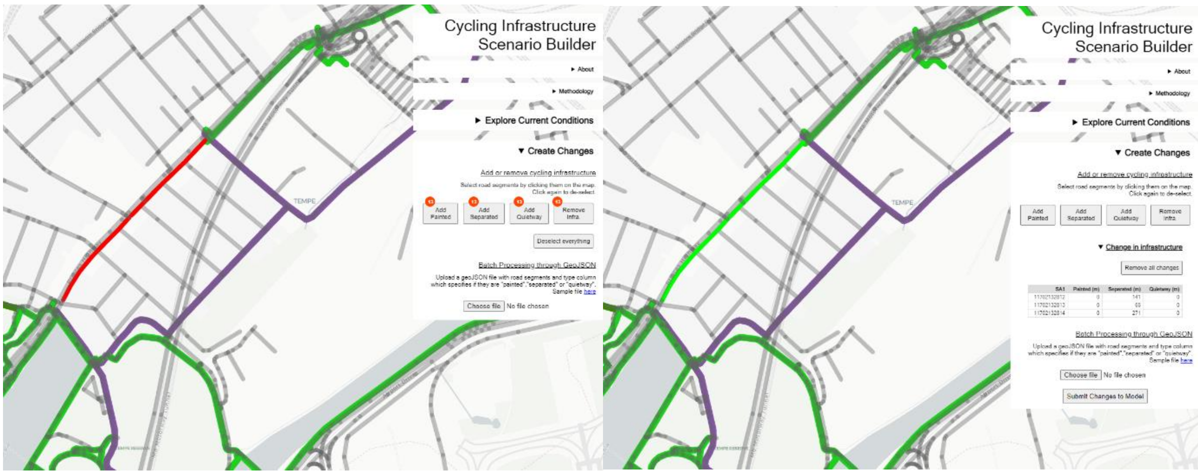

Figure 16. User interface for creating new cycling infrastructure scenario (in red)

Estimate uptake in cycling participation
The tool is able to take a user created cycling infrastructure scenario as input, and estimate the resulting cycling uptake. To produce estimates of cycling uptake, user input needs to include the shape, type, and location of the new cycling infrastructure. User input can be submitted to the model either through using the interactive web interface, or through uploading the infrastructure scenario in the batch processing mode.

Once an infrastructure scenario is submitted to the tool, the underlying model behind the tool will account for the type, length, and location of new cycling infrastructure. The effect of the new infrastructure is reflected in modelling, by changing individual attributes in the synthetic population; individuals adjacent to the new infrastructure would have access to more cycling infrastructure, and better connection to points-of-interest (POIs), which increases their likelihood to ride a bike. An important role of the model is in assessing and quantifying the extent of cycling behaviour change following new infrastructure investment.

Increase in cycling uptake is measured as:

- Increase in cycling participation
- Additional people cycling for transport purposes (from recreational cycling only)
- Additional people cycling for recreation/exercise (from utility-commuting cycling only)
- Additional number of transport related cycling trips per week
- Additional number of recreation/exercise cycling trips per week

Cycling uptake resulting from new cycling infrastructure is estimated as the difference in model predicted cycling participation, with or without the new infrastructure. The number of additional cycling trips are estimated based on cycling frequency of different types of cyclists. Modelling methods are discussed in more details in later sections.

Local attitudes towards cycling were collected through the cycling survey (described below in Section 2.1 Cycling survey), and its effect is reflected in all model predictions. Attitudes towards cycling is an important determinant of cycling participation. This attitudinal factor affects the “potential” of an area to develop cycling activities. Areas with highly favourable attitudes towards cycling but that lack of cycling infrastructure can be viewed as having a high potential for cycling, if appropriate infrastructure were provided. Conversely, areas with less favourable attitudes but good cycling infrastructure would need significant shift in cycling culture in order to have high levels of cycling participation.

The tool has a potential to explain and forecast changes in attitudes towards cycling resulting from either demographic shifts or behaviour change interventions and promotional campaigns. Australian cities are dynamic with fast-changing demographics and cultural make-ups. More favourable attitudes can result from various types of tangible or non-tangible intervention strategies. In order to reflect the would-be effect of having more favourable attitudes towards cycling, the tool incorporates a “cycling potential” estimation function, where cycling sentiment from the area with the most favourable attitudes towards cycling is transplanted to other parts of Sydney.

1.3 Interpretation of modelling results

For each hypothetical cycling infrastructure scenario, the dashboard measures the effect of new cycling infrastructure by estimating the additional number of people cycling, and the number of additional cycling trips. Outputs from this tool could have a wide range of possible uses and applications. In order for outputs from this tool to be applied correctly and in an appropriate context, in this section we note several aspects regarding the interpretation of outputs from the tool. Examples of interpretation of modelling results are given in *Section 3. Applications and Use Cases*

Cycling infrastructure scenarios with a higher predicted number of new cyclists and cycling trips are interpreted as being more effective than other scenarios with a lower predicted number. Such scenarios are often characterised by good quality cycling infrastructure in dense urban areas with relatively favourable local sentiment. Comparisons across different scenarios would therefore be useful for evaluating different proposals for new cycling infrastructure investment.

Different infrastructure scenarios may result in similar predicted numbers of cycling uptake. Under such circumstances, it is recommended that users consider other contextual information provided by the dashboard, such as terrain, and alignment with existing cycling infrastructure. Users should also consider the coverage-ridership trade-off which is also a major consideration, and has major equity implications. Concentrating cycling infrastructure in a small but densely populated area may encourage more local residents to ride a bike, but its effect will be limited spatially. Good spatial coverage of cycling infrastructure ensures equitable access by people living in different places. In summary, users should exercise discretion when using this dashboard, and the predicted cycling uptake should not be the only criteria for comparing different infrastructure scenarios, especially where there are similar estimates for cycling uptake.

The model's output predicts the change in the number of cyclists in a given area or community when a specific type of bicycle infrastructure is added. Therefore, the model has the potential to supplement the trip generation model in strategic transport modelling through appropriate modifications.

We recommend interpreting the current model's prediction of cycling participation as a type of "index" for comparison across different cycling infrastructure scenarios in terms of location, route, and type of cycling infrastructure. The project team does not recommend interpreting outputs from the tool literally as predictions for the number of cycling trips or the amount of cycling traffic on a specific roadway segment.

While cycling participation may fluctuate on a daily and temporal basis, the relative ranking of locations with higher cycling activities generally remains more consistent. Therefore, outputs from the tool will provide valuable insights in comparing different locations and varying infrastructure scenarios. Outputs from the tool reflect the level of cycling participation from local residents, which is different from prediction of cycling traffic that result from conventional transport planning models.

Relation to cyclist count/volume data
Outputs from this tool does not relate directly to the cycling traffic volume in a specific area or on a specific road segment. Cycling volume is often measured by either conventional traffic counters, or by crowd-sourced approaches such as through mobile apps such as Strava. Cycling volumes on a particular road segments depend on a wide range of factors, including the amount of cycling traffic generation and attraction, and the layout and function of the road network, which may channel or disperse cycling traffic depending on the topography of the road network. Accounting for route or areal level cycling traffic would therefore require estimates for both the generation and attraction of cycling traffic, and cycling route choice modelling, which is beyond the scope of this tool though they could be considered for potential future upgrades.

2. Data and Modelling Methods

This section discusses the data and modelling methods of the tool, and details how the under-the-hood model converts a user-provided cycling infrastructure scenario as input, and provides estimate for the resulting uptake in cycling participation as output.

Calibration data for the model is provided by a recent cycling survey, which collected information about cycling behaviour and tendencies of the Greater Sydney population. Modelling for the dashboard adopts a disaggregated individual level modelling approach in estimating cycling participation, meaning that the model predicts the likelihood to cycle for each “person” in the synthetic population (described below in Section 2.3 Synthetic population). Model prediction for each individual is then aggregated across all persons in a synthetic population[1] model to produce an estimate for overall cycling uptake. Two major modelling steps are involved in this process, namely, a discrete choice model, and a synthetic population model. The flow chart in Figure 17 shows the framework of data and modelling methods.

[1] The synthetic population model is a statistical representation, including demographics, access to cycling infrastructure and attitude toward cycling, of each adult resident in Greater Sydney. This is explained in more detail in 2.3 Synthetic population.

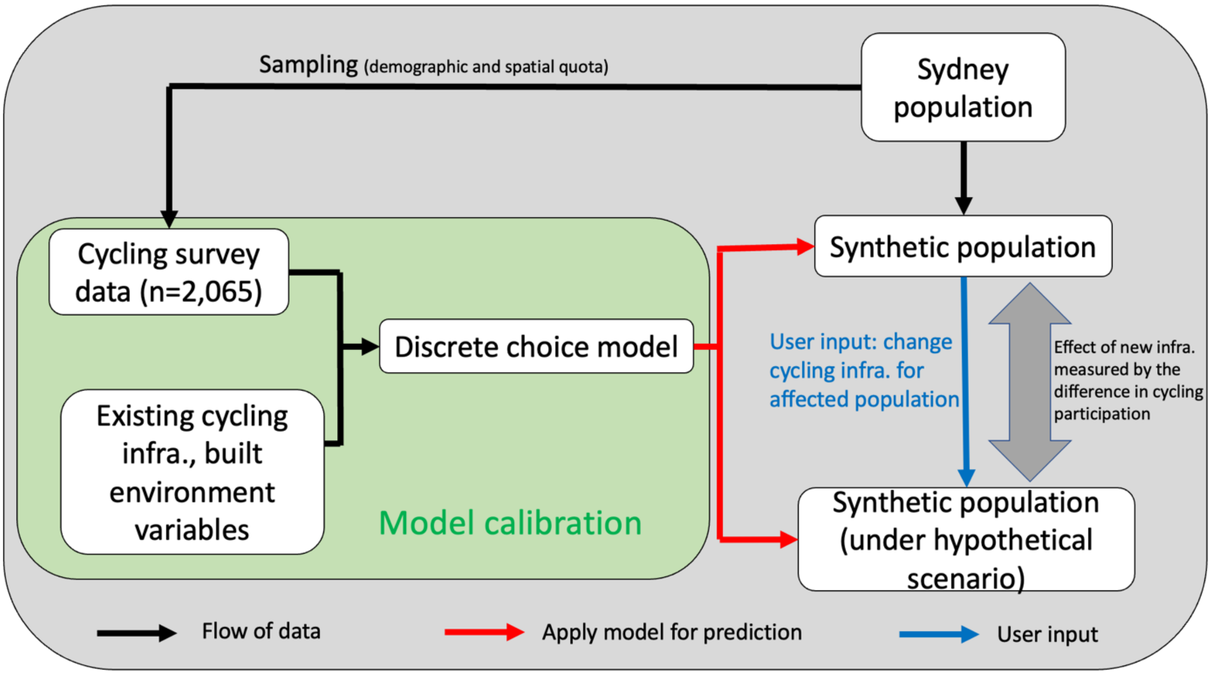

Figure 17. Flowchart showing the framework of data and modelling methods

2.1 Cycling survey

In order to gather data in support of the modelling work, the project team conducted a cycling survey of residents of the Greater Sydney area. The cycling survey has two major purposes, namely, to calibrate the cycling behaviour and tendencies of the Greater Sydney population (i.e. for the discrete choice model), and to help build the synthetic population model by supplying the local cycling sentiment variable (i.e. for the synthetic population model).

The survey collected detailed individual level data on social demographics, family structure, current cycling participation and other cycling behaviour, preferences and attitudes about cycling. Respondents' postcode at their residential location is also collected, which enables additional location-based factors to be used as explanatory variables for cycling behaviour.

This survey was conducted as an online survey with interlocking demographics and spatial quotas based on local government areas (LGAs) in order for the sample to be representative both demographically and spatially for the Greater Sydney population. The rationale for applying both spatial and demographic quota sampling is that people choose where they live, and how they travel as a single package (Zondag and Pieters, 2005), and people living in different areas may have different preferences and tendencies to cycle as a result of this self-selection; this interlocking demographics and spatial quota sampling method ensures that attributes from people living in different areas of Greater Sydney are captured.

The survey was conducted between October and December of 2022, and collected 2,065 valid responses. The survey was administered in collaboration with a market research company, Qualtrics, and by implementing an interlocking demographic and spatial quota sampling method to ensure the representativeness of the respondents. The survey was administered online, and the sampling frame consists of a double-opt-in market research panel that was managed by

Qualtrics. Data collected from the survey sheds light on the current cycling behaviour and tendencies of the Greater Sydney population at a detailed individual level, which provides the basis for calibrating the individual level discrete choice model for cycling participation. The survey data enables the effect of new cycling infrastructure to be modelled more accurately at an individual level. Attitudinal data collected from this survey also supported the construction of the synthetic population data (discussed in later sections), by providing information on how attitudes towards cycling vary across different areas of Sydney.

2.2 Discrete choice modelling

The discrete choice model predicts the likelihood of an individual to cycle, and is calibrated using individual level data from the cycling survey. The model establishes the connection between attributes of a person, characteristics of the person's residential location, and that person's likelihood to ride a bike. The model essentially extrapolates cycling behaviour patterns from survey respondents to all residents within Greater Sydney.

Logit (logistic) model is a discrete choice modelling approach that is commonly used for behavioural modelling. In the context of modelling cycling participation, logit models convert continuous individual level characteristics into discrete choice outcomes (cycle/not cycle). A previous literature review by the project team found that the likelihood to cycle for different purposes is driven by different factors, and therefore it would be more appropriate to model cycling separately for different purposes. We categorise cycling participation into different cycling purposes, including commute, utility, recreation and leisure cycling. In practice, we applied binomial logit models to estimate the likelihood of each person to cycle for each different purpose. Each person may participate in none, or any number of different cycling purposes. Equations (1) and (2) show the formulation of binomial logit models.

$$logit(cycling\ purpose) = \log\left(\frac{\mathrm{p}}{1-\mathrm{p}}\right) \qquad (1)$$

$$logit(purpose) = f(i) \qquad (2)$$

Where,

p: Likelihood of a person to cycle

$f(i)$: Linear combination of attributes of an individual (i)

The logit model is calibrated using cycling survey data. The dependent variable (dummy 0/1) is determined by whether a respondent has experience cycling for certain trip purposes in the past year. This definition of a cyclist differs from many other sources, such as the ABS census, which only identifies a person as a commute cyclist if that person cycled to work on census day (Australian Bureau of Statistics, 2021). By relaxing how recently a person needs to have cycled in order to be considered a cyclist, this definition of a "cyclist" captures both people who cycle regularly, and those who only cycle occasionally. The rationale for including both regular and occasional cyclists is to distinguish between those with any recent cycling experience, and those without, and to uncover fundamental differences that separate cyclists from non-cyclists.

Explanatory variables of the logit model include the attributes of that person, and characteristics of the built environment in that person's residential location. The list of explanatory variables is shown in table 2. The odds of a person to cycle (i.e. the natural log) is determined by a linear combination of explanatory variables. Higher odds correspond to greater likelihood to cycle for a particular cycling purpose. Variables that contribute to higher likelihood to cycle (higher odds) have a positive sign in the linear combination, and vice versa. This ease of interpretability is another benefit of logit models, i.e. explanatory variables are clearly distinguishable by whether they contribute to or impede cycling participation.

The modelling of this tool adopted an individual level behavioural model, which has a number of advantages over aggregate level models (e.g. direct demand models). First, modelling individual behaviour provides the opportunity to incorporate a wider range of factors that are sensitive to external interventions, such as the distance between home address and new cycling infrastructure, and individual level attitudes and preferences, which are not achievable through aggregate level modelling. Second, individual level modelling alleviates aggregation errors (such as Simpson's paradox) where a population level trend might disappear or even reverse when individuals or specific population groups are studied separately. This makes disaggregate level models more accurate and reliable. And lastly, modelling and predicting cycling participation at an individual level provides greater flexibility to study the effect of interventions on different population groups.

Since model predictions are made for each individual, the disaggregated modelling approach also provides greater flexibility for further aggregating model outputs based on different age, gender, and other population segments to examine the impact from interventions. Later sections provide a more detailed explanation for each step of the modelling approach. The discrete choice model is able to account for changes in cycling participation due to a variety of factors. The effect of changing demographics, personal attributes, or change in proximity to cycling infrastructure and POIs can be modelled by altering explanatory variables at an individual level. The overall change in cycling participation is obtained by aggregating changes across all individuals in an area.

2.3 Synthetic population

The synthetic population model is built using publicly available ABS census data, which includes representation for each adult in Greater Sydney. The population model provides the basis for applying the calibrated discrete choice model to predict cycling participation, and to aggregate model predicted cycling participation across all individuals living in the area to produce areal level estimates.

In order to apply the disaggregate model for cycling participation, a synthetic population model is built to represent the adult population of Greater Sydney at an individual level. This synthetic population model is a statistical representation for each adult in Greater Sydney, which comprises a large data table, with rows and columns. Each individual person is represented in this synthetic population model by a row; attributes of that person are described by respective columns. Table 2 below shows the basic structure of this synthetic population model.

Individual	**Age**	**HH Income category**	**Gender**	**Level of education**	**Home SA1**	**…**
1	55	below 800	M	Below bachelor	**12801160440**	…
2	32	above 3500	F	Postgrad	**11703133837**	…
3	19	800-1500	F	Bachelor	**12403146426**	…
…	…	…	…	…	…	…

Table 2. Hypothetical example of the synthetic population

The spatial resolution of the synthetic population data is at Statistical Areal level 1 (SA1). The calculation of the synthetic population model begins by obtaining the total number of residents at SA1 level by age and by gender from the ABS General Community Profile (GCP); a row is generated to represent each person based on their place of residence (SA1).

Other attributes in addition to age and gender are joined at an individual level using conditional probabilities. For instance, if x% of females aged 25 to 34 in a certain SA1 have a bachelor's degree, then we would assign a bachelor's degree to x% of females aged 25 to 34 in that SA1. We sequentially added other attributes probabilistically. Several location based variables, such as access to POI and to cycling infrastructure, are joined to the synthetic population based on synthetic individuals' residential SA1 location.

In addition to location based (SA1) social demographics and built environment data, we added attitudinal variables probabilistically to individuals in the synthetic population data, based on where an individual lives (based on LGA). We categorize attitudes towards cycling sequentially into 4 groups; during the cycling survey it was observed that certain areas of Greater Sydney are associated with more individuals having more favourable attitudes towards cycling than in other areas. In order to reflect this geographical variation in attitudes towards cycling, we probabilistically assign attitudinal classifications to individuals in the synthetic population with the same proportional as the survey data based on geographical locations (LGA). Table 3 below shows the list of attributes contained in this synthetic population.

Type	Attribute name
Social demographics	Age
	Gender
	Live with child(ren)
	Live with a partner
	HH Income
	Level of education (below bachelor, bachelor, postgrad)
Residential location related	Access to cycling infrastructure
	Access to points of interest (POI) (services, parks)
Attitudes	Attitude towards cycling (categorical, 4 types)

Table 3. Attributes (explanatory variables) of the synthetic population

2.4 Accounting for the effect of new cycling infrastructure

Both the amount of cycling infrastructure, and the number of POIs that are adjacent to an individual's residential location are included as explanatory variables in modelling cycling participation. Dashboard users are able to create new cycling infrastructure through the interactive user interface. Newly added cycling infrastructure would have the following effect on nearby residents:

- Increase the amount of available cycling infrastructure
- Increase connectivity to POIs

Once the new cycling infrastructure scenario is submitted to the model, a catchment area is generated to identify residents that are affected by the new cycling infrastructure. If one segment of newly added cycling infrastructure is within threshold distance of a nearby residents, then the entire length of that segment will be added to the amount of cycling infrastructure already accessible to that individual (the average length of a road segment is around 30 metres). In order to reflect the actual access to new cycling infrastructure, the distance threshold between residents and cycling infrastructure is based on road network distance instead of Euclidean distance; therefore, new cycling infrastructure will not be considered as accessible to residents in cases where natural or man-made barriers exist and the road network distance between the resident and cycling infrastructure exceeds the threshold.

Effect of new cycling infrastructure

A focus in the modelling is to examine the effects of cycling infrastructure on cycling participation. Cycling infrastructure takes various shapes and forms. In order to accurately model the effect of cycling infrastructure, we categorise existing cycling infrastructure into 3 types, namely,

a) Separated paths where cyclists have exclusive right-of-way, do not mix with pedestrians, and are physically separated from motor vehicle traffic;

b) Other cycling infrastructure, where signs and markings allow cycling on footpaths or road shoulders; and

c) Quietways, where low traffic volume and/or speed limit conditions are present, or where cycling is permitted on a one-way street.

Conditions on quietways can vary noticeably depending on location, and logit models could not accurately capture the effect from quietways on cycling participation. Therefore, we did not include quietways during model calibration. However, quietways is increasingly becoming an important component in cycling infrastructure in Greater Sydney, and there is a need to understand their effect on cycling participation. Based on survey responses, quietways are measured by respondents to be around 85% as comfortable as separated bicycle paths. Therefore, in modelling, new quietways are modelled to have 85% of the effect as separated cycling path of the same length. Table 4 shows the classification of cycling infrastructure types.

Infrastructure category	**Separated paths**	**Other cycling infra, painted**	**Quietway**
	Bicycle Paths	Shared Use	Quiet Streets
		Bicycle Lane	Contra-flow Cycling
		Bicycle Only	

Table 4. Classification for cycling infrastructure types

The model includes the amount (km) of different cycling infrastructure in close proximity of each person as an explanatory variable in explaining their current cycling participation. A person's likelihood to cycle is increased with additional cycling infrastructure in close proximity to that person's residential location; well-placed cycling infrastructure that is close to more residents increases access to cycling infrastructure to more persons, which is likely to produce a higher uptake in cycling participation.

In presenting the effect of new cycling infrastructure in model output, the dashboard shows the increase in overall cycling participation, and the number of people who would start cycling for transport or recreation/exercise purposes who did not cycle for such purposes previously. The dashboard also shows the estimated number of additional transport and recreation/exercise cycling trips, based on the cycling frequency of different types of cyclists, which is obtained from the cycling survey.

Effect of cycling infrastructure in improving access to POIs
In addition to cycling infrastructure, cycling participation is also sensitive to the number of POIs connected by new cycling infrastructure. New cycling infrastructure that is located near a POI and also in close proximity to a resident has greater potential to be used by that resident in accessing the POI. The effect of such cycling infrastructure is reflected in a greater number of POIs reachable by cycling.

To reflect the effect of cycling infrastructure in increasing access to POIs, catchment areas are generated from new cycling infrastructure locations and from POIs; residents that live within the combined catchment areas are assumed to have enhanced access to these POIs. At an individual level, being able to reach a greater number of POIs (as a result of new cycling infrastructure) increases the likelihood to cycle; this effect is quantified by models in evaluating the effect of new cycling infrastructure.

Effect of attitudinal change (cycling potential)
The cycling survey conducted by the project team identified more favourable attitudes towards cycling in some LGAs than in others. On average, residents in the City of Sydney have the highest levels of positive attitudes towards cycling. Since the attitude towards cycling has a notable effect on cycling participation, a general cultural shift and change in attitudes towards cycling will affect cycling participation. This cultural shift or change in attitude can be created via a range of strategies, including public information campaigns, rider training programs, incentives, and events or other promotions. In order to account for this possible change in cycling culture due to non-infrastructure factors, this tool includes a “cycling potential” scenario, where the composition of attitudes towards cycling from the City of Sydney is “transplanted” to other LGAs.

3. Applications and Use Cases

In this section we showcase potential applications of this tool. We demonstrate the use of the *Cycling Infrastructure Scenario Builder* in assisting with both state and local level cycling infrastructure planning, namely in the planning of strategic cycling network, and in project evaluation and prioritisation of local cycling infrastructure projects.

Demonstration of use cases focused primarily on the predicted cycling uptake from new cycling infrastructure. Users in actual use cases should consider a wider range of factors in comparing scenarios and prioritising cycling infrastructure investment, and the predicted cycling uptake should not be the only criteria, especially when the differences in cycling uptake were small.

3.1 Strategic cycling network planning

Strategic cycling networks provide connection over greater distances than local cycling routes. Strategic cycling networks are usually created as part of long-term strategies for connecting places, and the establishment of a strategic cycling network can help safeguard critical corridors against competing land uses (Bicycle NSW, 2023).

In the use case demonstration, we show an example of comparing two cycling infrastructure scenarios in the dashboard. The first scenario includes a hypothetical cycling path connecting Chatswood and Lindfield via Pacific Highway (Figure 18); the second scenario includes a hypothetical cycling path connecting St Leonards and Lane Cove via Berry Road, River Road, Northwood Road and Longueville Road (Figure 19). Both scenarios include similar length of additional cycling infrastructure (3.4 km and 3.5 km, respectively).

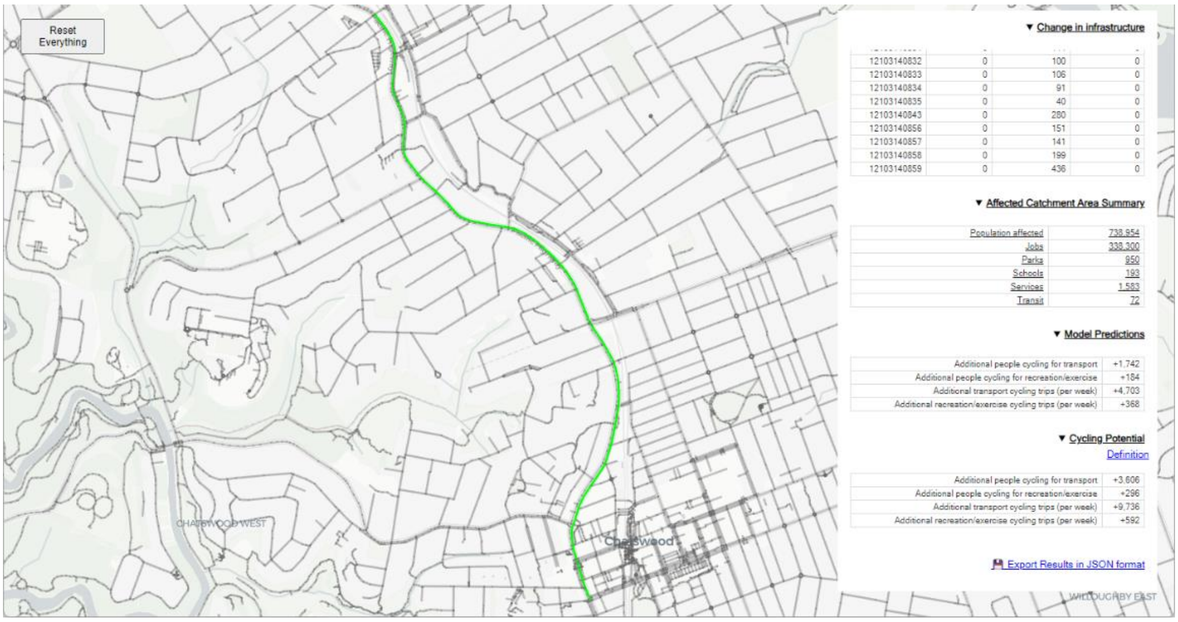

Figure 18. Hypothetical cycling path connecting Chatswood and Lindfield via the Pacific Highway, (Approx. 3.4 km, estimated 1,926 new cyclists)

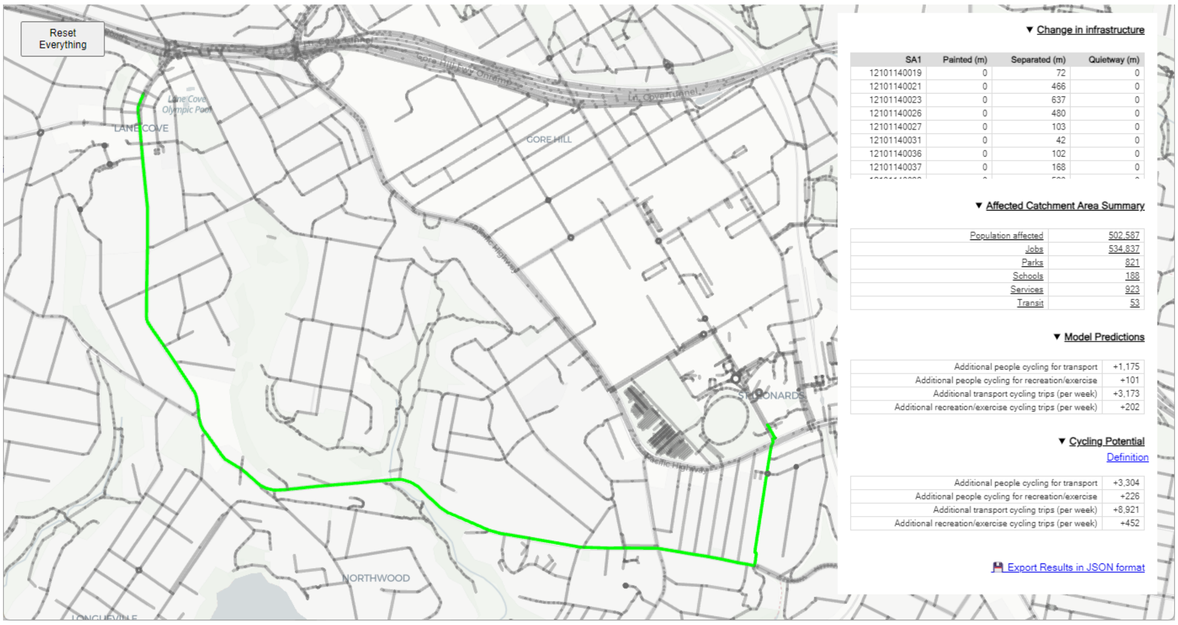

Figure 19. Hypothetical cycling path connection from Lane Cove to St Leonards via Longueville Road, Northwood Road, River Road and Berry Road (Approx. 3.5 km, estimated 1,276 new cyclists)

These two scenarios are submitted to the web interface to evaluate the effectiveness of new infrastructure in terms of generating new cyclists. The Chatswood – Lindfield scenario is estimated to produce 1,926 new cyclists and 5,071 new cycling trips per year. This is significantly higher cycling uptake than the Lane Cove – St. Leonards scenario, with a predicted 1,276 new cyclists and 3,375 new trips, despite the similar distance and a higher population densities around the Lane Cove – St Leonards corridor. This suggests that the Chatswood – Killara route would be a more effective investment within a strategic cycling network planning framework. However, the cycling potential between the two routes are much more similar (3,902 new cyclists and 10,328 new trips for Chatswood – Killara vs 3,530 new cyclists and 9,373 new trips for Lane Cove to St Leonards). This highlights the impact of the cycling sentiment index on predicted ridership – Figure 20 shows the different level of positive feeling about cycling in the area of these two hypothetical routes.

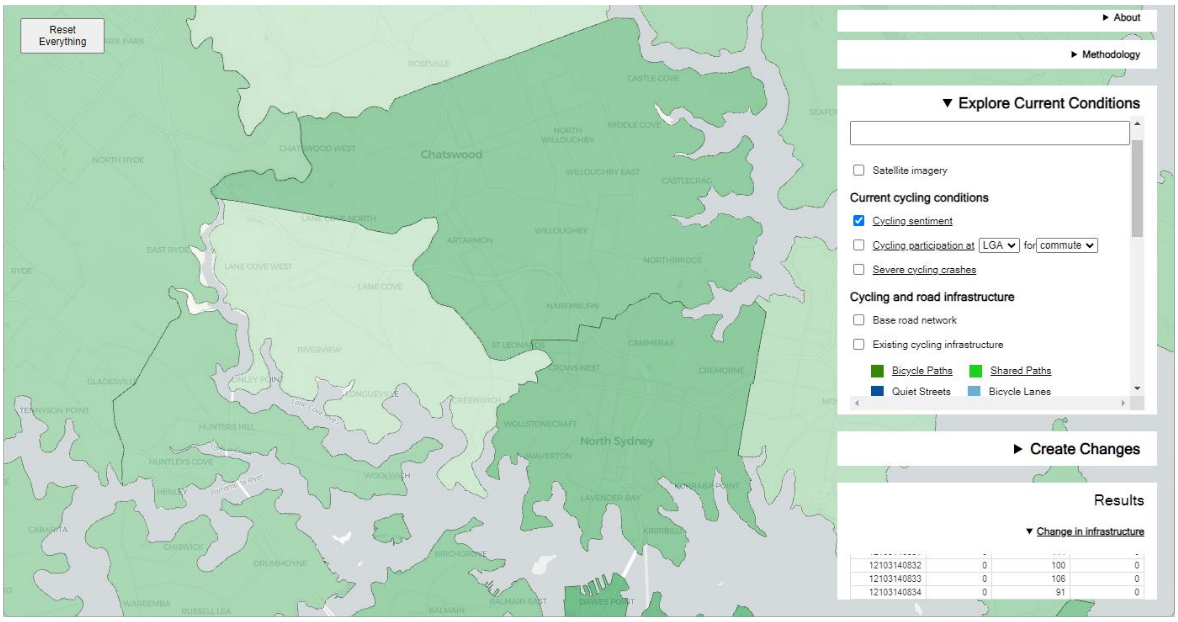

Figure 20. Cycling sentiment in North Sydney

3.2 Local cycling infrastructure project evaluation and prioritisation

The tool can also be applied to evaluate and compare different local cycling infrastructure investment plans and strategies. In this section we showcase an example of using the tool to evaluate and prioritise cycling infrastructure in a local area.

The example scenario involves building a hypothetical cycleway along Alison Road in Randwick, with an extension either northward along Avoca Street (Alison Road – North scenario, Figure 21), or an extension southwards along Avoca Street (Alison Road – South scenario, Figure 22). Both scenarios include a similar length of additional cycling infrastructure (Approximately 1.1 km).

Figure 21. Hypothetical cycling path, Alison Road/Avoca Street – North scenario, (Approx. 1.1 km, estimated 1,379 new cyclists)

Figure 22. Hypothetical cycling path, Alison Road/Avoca Street – South scenario, (Approx. 1.1 km, estimated 1,321 new cyclists)

The bikeability scenario builder provides a very similar estimate for the two scenarios (1,379 new cyclists in the North scenario vs 1,321 new cyclists in the South scenario), meaning that other considerations, including connections to other existing or planned infrastructure, priority POIs or schools, or identified areas with a history of crashes, should take priority when choosing a route. In this case, the North scenario would provide better connections to local schools and address prior crash locations, while the South scenario would provide better connections to a local shopping centre (Figure 23).

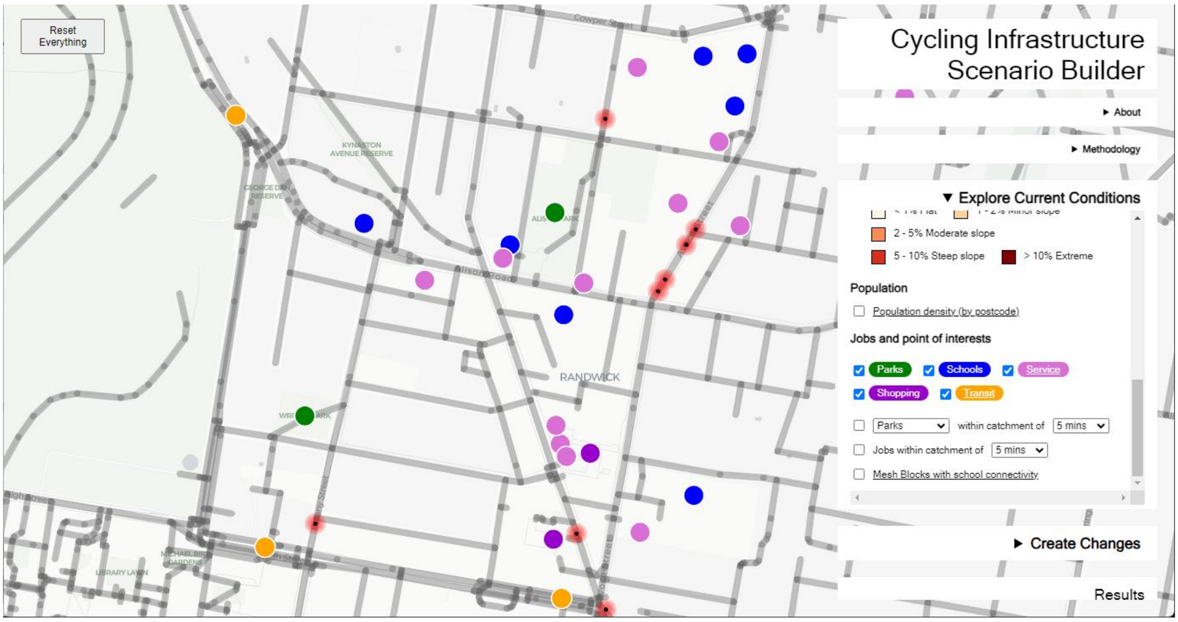

Figure 23. POIs and serious bicycle crash sites near the two scenario routes

The tool is designed to be a valuable resource for planners and policy makers by assisting in effective, appropriate and cost-efficient provision of bicycle infrastructure. Improving cycling environments are critical to increasing cyclists riding for commute, utility transport, recreation and exercise. There are important individual and societal benefits from riding a bicycle. Riding is good exercise and can contribute to individual physical and mental health, social inclusion and wellbeing. When an individual chooses to ride a bike instead of driving or riding in a motor vehicle, there are environmental benefits from reduced carbon and pollution emissions, and reduced demand for road space. In comparison with most other modes of transport, riding a bicycle is inexpensive for the individual and in terms of cost of infrastructure provision. This tool provides data-driven support for planners and policy makers as they decide where to build new infrastructure to maximise benefits for new and existing cyclists. In this way, this tool can help contribute to a more sustainable city.

4. Future Research and Limitation

The model's output predicts the change in the number of cyclists in a given area when a specific type of cycling infrastructure is added. The model has the potential to supplement the trip generation model in strategic transport modelling through appropriate modifications. Trip generation models typically predict the amount of travel activity by considering land use characteristics, population characteristics, and built environment factors. Most current models are limited to forecasting demand for automobiles and pedestrians, thus unable to predict cycling demand or account for changes in cycling infrastructure.

The project team does not recommend interpreting outputs from the tool literally as predictions for the number of cycling trips or the amount of cycling traffic at the street level. To predict cycling traffic volume at the street level during specific time periods, such as peak hours, it is essential to accurately forecast travellers' mode choice and route selection. In Sydney and Australia, cycling is predominantly used for recreational and leisure purposes rather than commuting. Applying conventional mode choice models directly poses challenges because their fundamental principles are based on the idea that mode selection is determined to minimise travel time and cost.

Another reason for this recommendation is the inherent inaccuracy in estimating cycling frequency based on stated behaviour survey. There is also significant temporal and seasonal variation in the level of cycling participation, and there is no reliable method yet to estimate the number of cycling trips within a particular time period.

5. Acknowledgements

The project is funded by Australian Research Council grant (LP200201073). The cycling survey was funded by the UNSW Digital Grid Futures Institute (DGFI) funding (PS66643). We would like to acknowledge the significant contribution from our project partners from NSW Office of Sport, Penrith City Council, Transport for NSW, Wollongong 2022 Limited, and Wollongong City Council for providing continuous support throughout the course of the project. Both the practicality and usability of tool are greatly improved by suggestions and feedback received from our project partners.

References

Australian Bureau of Statistics. General community profile (WPP), (2021).

Bicycle NSW, “Strategic cycleway corridors explained”, (2023), URL: https://bicyclensw.org.au/strategic-cycleway-corridors-explained/#:~:text=The%20Strategic%20Cycleway%20Corridors%20are,of%20new%20and%20improved%20routes.

Munro, C. "National walking and cycling participation survey" Technical report, Cycling and Walking Australia & New Zealand, (2023).

Zondag, Barry, and Marits Pieters. "Influence of accessibility on residential location choice." *Transportation Research Record* 1902, no. 1 (2005): 63-70.