



APPENDIX B

Guidance for Utilizing GIS with the ActiveTrans Priority Tool

This appendix addresses two questions associated with using GIS to implement the ActiveTrans Priority Tool (APT) methodology.

1. How can GIS complement the APT?
2. What are some key considerations associated with using GIS to complement the APT?

How Can GIS Complement the APT?

GIS can complement the APT in the following ways:

- To spatially define the improvement locations (Step 8).
- To measure variables that include spatial relationships (Step 8).
- To represent prioritization outcomes spatially so that they can be more easily reviewed and communicated to the public and other stakeholders (Step 10).

GIS can be used for other aspects of the method, including scaling variables (Step 9), applying factor weights (Step 3), and calculating prioritization scores (Step 10); however, doing so may compromise the transparency of the prioritization process. While GIS systems can typically perform the calculations needed for scaling, factor weighting, and calculating prioritization scores, they usually do not retain a readily accessible record of the formulas used, making it difficult to check or adjust the calculations after the fact. Consequently, it is recommended that the scaling, weighting, and prioritization calculations be performed in a spreadsheet tool, which is capable of retaining a record of formulas and can be accessed and reviewed by people who are unfamiliar with GIS or do not have it on their computers.

What Are Some Key Considerations Associated with Using GIS to Complement the APT Method?

A variety of issues must be considered when using GIS to implement aspects of the APT. Key considerations are discussed below:

- Choosing a clear process.
- Documenting the process.
- Checking data for accuracy.
- Defining the improvement location extents.
- Setting up the GIS database.
- Selecting buffer sizes.
- Normalizing calculations.
- Double-checking calculations.
- Communicating results.

Choosing a Clear Process

Any calculations performed in GIS will likely need to be communicated to people who are unfamiliar with GIS. As a result, it is probably better to opt for simpler GIS procedures that can be more easily communicated rather than more complex procedures that may be more difficult to explain and understand. For example, there might be a legitimate case for using raster analysis for some prioritization calculations; however, the potential benefits of this type of analysis must be weighed against the difficulty of explaining it.

Documenting the Process

It is important to keep a thorough, written record of all calculations performed in GIS as part of the APT methodology. For example, if buffers are used to calculate a particular variable, it is important to record the buffer size and the rationale for selecting it. If raw variable values are normalized, then details regarding the normalization should be recorded, and so on.

Keeping a detailed, written record is critical for transparency, so that others can understand how the calculations were derived, and is also valuable for reproducing the process, which may be necessary for a variety of reasons. For example, it is not uncommon for additional improvement locations to be added to a prioritization process after the calculations for an initial set of prioritization locations are complete. Calculations for the additional improvement locations should be handled in the same way as the initial calculations for consistency, in which case a thorough written record can be extremely helpful. Also, if an agency has established a process for prioritizing improvement locations, and that process is regarded as successful, it will likely want to follow the same procedure in the future. Having a written record of the process, including detailed GIS notes, will help the agency duplicate it even in cases where there have been personnel changes.

Checking Data for Accuracy

It is important to double-check the accuracy of any datasets that will be the basis for GIS calculations before the GIS calculations are performed. Datasets that do not accurately reflect the intended conditions may result in significant calculation errors. For example, if proximity to schools is selected as a variable but the dataset used does not include private schools or schools built in recent years, then improvement locations near schools not reflected in the data may receive lower prioritization scores by mistake. Another example may be crash data, which may have been transcribed from hard copy reports. It is often good practice to conduct an accuracy scan by randomly selecting data points and comparing them to the reports. If inaccuracies are uncovered, the agency must consider appropriate next steps which, depending on the nature and extent of the inaccuracies, might include correcting the inaccuracies; using a different, more accurate dataset; or dropping variables if accurate data cannot be found to represent them.

Defining the Geographic Extents of Improvement Locations to Facilitate Comparison

Defining the geographic extents of improvement locations is something that will likely be done as part of Step 7: Set Up Prioritization Tool. GIS can be used to define the geographic extents of improvement locations so that they are more comparable, something that may be particularly important in the case of roadway segments, corridors, and neighborhoods/areas, since in these cases geographic extent can have a significant impact on the final prioritization score. It is not necessary that geographic extents be exactly equal, only that they be roughly equivalent, avoiding extreme differences that can result in counterintuitive prioritization results.

Setting Up the GIS Database

Setting up the GIS database is something that should be accomplished in Step 7: Set Up Prioritization Tool. As previously mentioned, in order to maintain transparency in the prioritization process GIS should generally not be used for scaling variables, applying factor weights, and calculating prioritization scores. These calculations should instead be done in a spreadsheet tool. Consequently, it is recommended that the GIS database be set up to facilitate conversion between spreadsheet and GIS database formats easily. The GIS database should generally include the following columns or fields:

- Common key field. A common key is an attribute that two data sets have in common, such as a street segment ID. Using a common key ensures that the data order and integrity is maintained as several sources are combined.
- Fields for important improvement location descriptors. Needed descriptors will depend on the prioritization purpose and future filtering needs. For example, an agency may wish to filter improvement locations by street or neighborhood, in which case fields should be added for the street and neighborhood names.
- Fields for raw variable values. These are the fields that will accommodate the unscaled value of each variable after the necessary spatial calculations have been performed.
- Fields for scaled variable values. The purpose of these fields would be to facilitate transfer of scaled variable values calculated in a spreadsheet tool for display on a map.
- Fields for weighted factor (or variable) values. The purpose of these fields would be to facilitate transfer of weighted factor (or variable) values calculated in a spreadsheet tool for display on a map.

Selecting Buffer Sizes

The selection of buffer sizes will likely occur in conjunction with Step 8: Measure and Input Data. Buffers are commonly used in GIS prioritization calculations; however, it is important to select buffer sizes carefully, since the choice of buffer size can have a significant impact on prioritization scores.

Figure B-1 and Figure B-2 show how the choice of buffer size can make a difference. Both figures show corridors X and Y and both figures show a series of stars representing locations for the variable. In Figure B-1 the buffers around Corridor X and Corridor Y are comparatively small, and both buffers contain four stars. In Figure B-2 the buffer is larger. As a result, the buffer around Corridor X includes 11 stars, while the buffer around Corridor Y includes only five stars, meaning that the raw score for Corridor X is now more than twice as large as the raw score for Corridor Y.

Given the impact of buffer sizes, it is very important to be clear about what the buffer is intended to capture. If the improvement location type is intersections and crossings, and the variable being calculated is “number of crashes involving pedestrians,” then the buffer is likely intended to gauge the relative current safety of the intersection or crossings. The choice of buffer size should reflect this intention, i.e., it should only include crashes at or very near the intersection or crossing rather than crashes that are further afield, since such crashes are unlikely to suggest anything about the relative safety of the intersections or crossings being prioritized.

Similarly, if the improvement location type is a corridor and the variable being calculated is “proximity to park,” then the buffer is likely intended to gauge the relative potential demand for bicycling along each corridor as a means of accessing parks. In cases like this, it might be tempting to consider how far people are generally willing to bicycle to parks, but this would

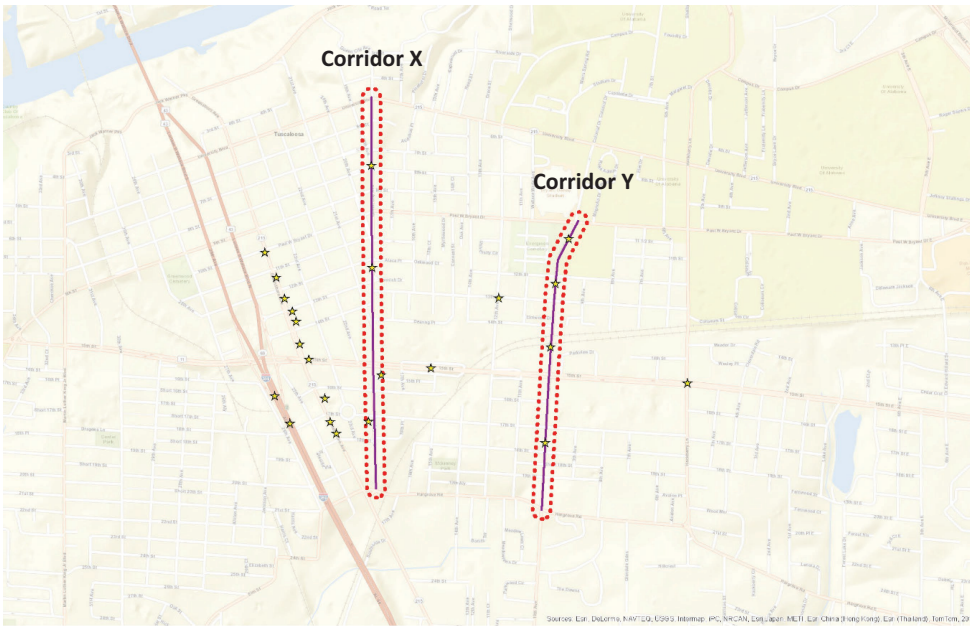


Figure B-1. Corridor X and Corridor Y with small buffer.

reflect a mistaken understanding of what the buffer is intended to capture. The question in this case is not, “How far would the average person be willing to ride a bicycle to a park?” but rather, “Would the average person use this corridor to access a given park if the park were X distance away from the corridor?” Thus, the way the question is framed has a significant impact on the chosen buffer size, with the former question seeming to justify a much larger buffer than the latter question.

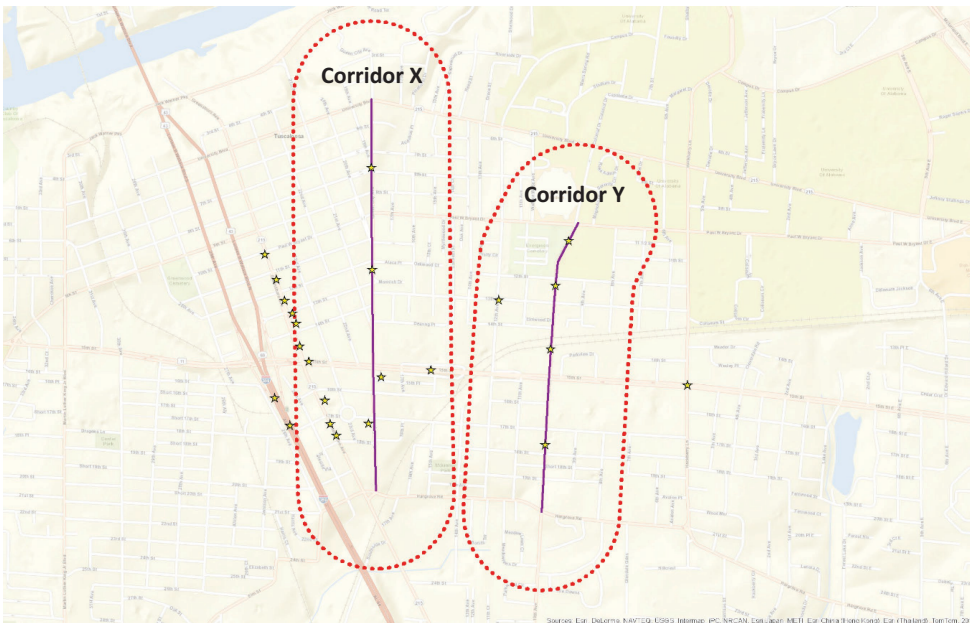


Figure B-2. Corridor X and Corridor Y with larger buffer.

In addition to understanding what the buffer is intended to capture, it is also important to consider whether the variable layer used in the calculations is represented by points, lines, or polygons, since the type of geographic representation can affect whether particular locations are included in the buffer. For example, it makes a difference whether parks are represented as points or polygons, since although the boundaries of a park may be within the buffer distance, a point representing the location of the park may not be within the buffer.

Finally, given the sensitivity of the buffer size choice and the number of factors that must be considered, it is recommended that buffer sizes be confirmed with others involved in the prioritization process before being used in GIS calculations.

Normalizing Calculations

The process of normalizing calculations is typically part of Step 8: Measure and Input Data. In many cases raw variable values involving roadway segments, roadway corridors, and neighborhoods/areas should be normalized in GIS before being transferred to a spreadsheet for scaling and other prioritization calculations. The reason is that the relative size of these location types will likely have an unintended influence on the raw variable value. For example, a larger neighborhood might have a larger number of community facilities simply because it is larger. Should it receive a higher prioritization score on this account alone? Normalization removes size from the equation. In the case of segments and corridors, this will likely mean dividing the segment or corridor variable raw value by the length of the segment or corridor in feet or miles, while in the case of neighborhoods/areas the divisor will likely be the area in acres or square miles. Table B-1 provides an example of how normalization works. In this example, the normalized variable value is calculated by dividing raw variable value for the corridor by the length of the corridor in miles.

Double-Checking All Calculations

All GIS calculations should be double-checked to make sure that they were performed correctly. Strategies for double-checking GIS calculations related to the APT include:

- Look at the high and low values for each variable. Are they unexpectedly high or low?
- Use an improvement location you know well as a test case. Do the values for this improvement location make sense?
- For each variable, create a heat map to symbolize the values. Do you notice anything unusual or unexpected when the data is displayed this way?

Table B-1. Example of normalizing variable values by corridor length.

Raw Variable Value	Corridor Length	Normalized Variable Value
30	1.5 miles	20
20	2 miles	10
45	3 miles	15

If the answer to any of these questions is yes, then the reasons why should be investigated, and if necessary the operations should be re-performed.

Communicating Results

Maps created in GIS are an excellent way to communicate and cross-check the final prioritization calculations, including scaled variable values, unweighted factor scores, weighted factor scores, and prioritization scores. Since it is recommended that these calculations be done in a spreadsheet tool rather than in GIS, they will likely have to be imported back into GIS in order to be mapped.

Several types of maps can be created. Examples include heat maps (Figure B-3) and phasing maps (Figure B-4).

Heat Maps

Heat maps can be used to display the relative magnitude of scaled variable values, unweighted factor scores, weighted factor scores, and prioritization scores for public review. They can also serve as the basis for further prioritization efforts, such as the identification of priority pedestrian or bicycle routes within the study area by a stakeholder group.

Phasing Maps

One result of a prioritization process may be to determine project phasing. In such cases, a phasing map can be used to communicate the prioritization process. A phasing map shows project routes or spot locations using a symbolic scheme to demonstrate a time frame for improvements. The example below shows short-, medium-, and long-term prioritized projects for the Bellingham, WA, Bicycle Master Plan.

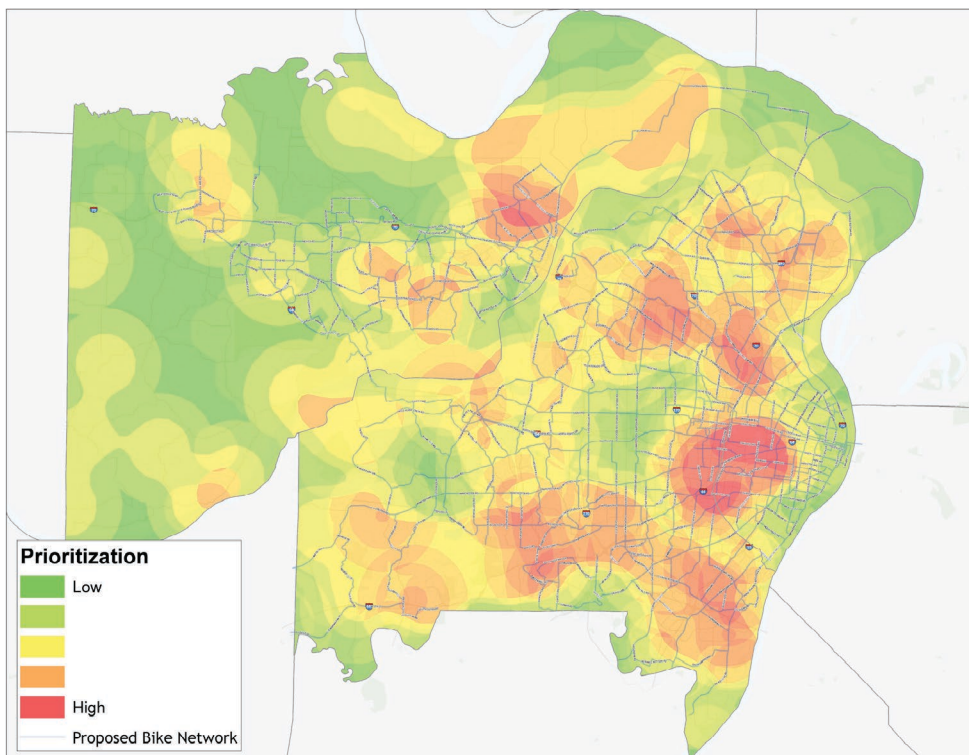


Figure B-3. Heat map.

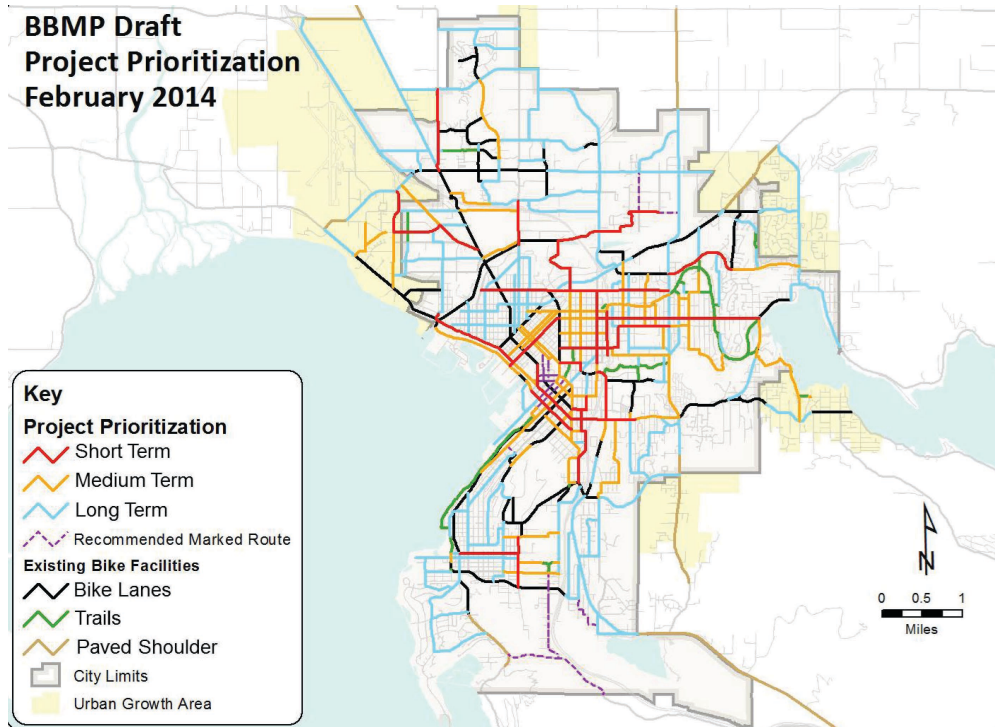


Figure B-4. Phasing map.