USING ARCHIVED TRANSIT DATA TO CONDUCT A TIME-SPACE ANALYSIS OF SAN LUIS OBISPO TRANSIT ROUTE 6b

Kevin Carstens, EIT

Phone: 916-521-8532

Email: krcarste@calpoly.edu

District: Western

Address:

5016 Blackstone Dr.

Rocklin, CA 95765

ABSTRACT

In today's fast-paced world, waiting for the bus can be agonizing, especially for a rider in a rush. Therefore, an in-depth analysis of why buses are delayed and perhaps how to solve such delays can be of enormous benefit. One such analysis was performed on San Luis Obispo (SLO) Transit's Route 6b, with service to downtown SLO, California Polytechnic State University's campus, and points in-between. Automatic vehicle location (AVL) data was collected and processed for buses traveling on this route during two weeks in the spring of 2015. This processed information yields valuable insight into the performance of Route 6b.

With the dwell and travel portions of the data separated, it is possible to assess dwell time and travel speeds for different segments of the route, and how they compare to the schedule. With this assessment, time can be shifted from areas that run ahead of schedule to areas that fall behind. Specifically for Route 6b, it was found that the bus traveled slower than the schedule allowed from the Kennedy Library stop to the Downtown Transit Center, and faster in the reverse direction. Likewise, on average the bus dwelled longer than scheduled at the Library, but less than scheduled Downtown. Therefore, it is recommended that the schedule for Route 6b be adjusted to account for these results, giving more time for the Library-to-Downtown segment and for dwelling at the Library, and less time for the Downtown-to-Library segment and for dwelling Downtown. This time shift requires no capital investment nor major route changes, yet increases bus on-time performance and service reliability. This kind of shift will also improve passenger perception and goodwill among the riding population.

The focus of this paper is on the methodology of ascertaining such conclusions from raw AVL data, with SLO Transit's Route 6b serving as an overarching example of its implementation.

TABLE OF CONTENTS

Introduction 1
Objectives 1
Data Analysis 2
Conclusions and Recommendations 8

Appendix A: Route 6b Schedule 10

TABLE OF FIGURES

Figure 1: Route 6b 3
Figure 2: Time-Space Diagram Depicting Runs vs. Real Time 5
Figure 3: Time-Space Diagram Depicting Runs vs. Elapsed Time 5

TABLE OF TABLES

Table 1: Average Segment Length, in Miles 4
Table 2: Proposed Dwell and Travel Time Changes (in Minutes) 8
Table 3: Proposed Schedule Changes 9

INTRODUCTION

As a frequent rider of San Luis Obispo (SLO) Transit's Route 6b, I have noted much room for improvement. Particularly, certain segments of the route are given too much schedule time and certain parts are given too little schedule time. Therefore, I sought to analyze the efficiency of this route with a goal of finding improvements to recommend.

ITS technology was the backbone of my analysis, in the form of automatic vehicle location (AVL) data. After acquiring and culling down data, I generated equations for the dwell and travel segments of the route. This was in keeping with my original objective to generate these equations. However, I realized that the equations were not very useful in helping determine ways to improve Route 6b. Therefore, I continued my analysis to determine dwell times and travel speeds, and compared them to the schedule. Through this comparison, I was able to generate some recommendations for changes to the schedule to better reflect how Route 6b is operating in reality.

OBJECTIVES

The overarching objective with this analysis was to take an in-depth look at SLO Transit's Route 6b. I had the AVL data, now I needed to convert it to something meaningful. I set out with the initial objectives of compiling this data into a time-space diagram depicting the bus completing each run across the day and versus the schedule, and determining equations for travel speed and dwell time.

However, after completing this analysis, I discovered that these objectives left much to be desired in regards to usefulness. Therefore, I amended my objectives to focus on comparing dwell times and travel speeds to the established schedule, rather than generating equations. With these

new objectives in mind, my end-goal was now to create a proposed new schedule that would better reflect real-world operating conditions. This new schedule will allow the bus to run on time more often and therefore become a more reliable and attractive option.

DATA ANALYSIS

The first step in this analysis was to acquire the raw data. For this, I turned to Bishop Peak Technology, the creators of the SLO Bus Tracker mobile application. They track bus AVL data, so I asked them for a snippet of data in which to analyze. What they sent back included data for all routes, from March 31st, 2015, to April 15th, 2015. Altogether, there were 333,323 lines of data. Since I was seeking a more focused analysis, I isolated just Route 6b, just April 1st, 2015 (a Wednesday), and just the daytime runs (no evening runs). This pared the data down to just 2,306 lines, a much more reasonable number.

With the data now reduced to only pertinent information, I began the data analysis. The first step was to convert the latitude and longitude data into distance. I utilized the "Great Circle" method, which approximates the Earth as a sphere (the same way that GPS satellites log data), then calculates the distance between two points on that sphere. The equation for this conversion is provided below, with an additional factor of 1.15078 to convert nautical miles to regular miles:

```
=[ACOS(SIN(RADIANS(lat2))*SIN(RADIANS(lat1))+COS(RADIANS(lat2))*COS(RA
DIANS(lat1))*COS(RADIANS(long2-long1)))*180/PI()*60]*1.15078
```

"lat1" and "lat2" represent the latitudes of two points, and "long1" and "long2" represent the longitudes of two points. This equation computed the spherical distance between these two points, which was then converted into miles.

Next, I needed to convert Coordinated Universal Time (UTC) to Pacific Daylight Time (PDT), a difference of 7 hours. After making this conversion, I then needed to codify the 15 second resolution of the timestamps. Before, each minute was repeated four times, once for each 15 second data point, so I simply needed to append increments of 15 seconds onto each timestamp.

With the data, I was not provided stop names, but instead a column of "NextStop" data, which had a numerical code signaling the next timed stop on the bus's path. After plugging coordinates into Google Maps and checking versus the route, I determined that "44" correlated to the Library to Downtown segment, "29" correlated to Downtown to Taft, and "0" correlated to Taft back to the Library. A map of Route 6b is provided in Figure 1, with these locations annotated.

After determining which segments correlated to which numbers, I was then able to establish a starting point for each loop. I chose the point where the data switches from "0" to "44" (when the bus arrives at the Kennedy Library stop) as the starting point. From there, I summed up the total distance traveled for each of the three segments. However, I soon found that due to aberrations in the GPS data (rounding, bad readings, and other systematic errors), the runs were different lengths. For example, for the total distance, the bus would take anywhere from 3.7 to 4.4 miles to complete the same route. I found the average for each of the three segments, then standardized all of the runs to these averages by creating a percent complete column, and finally multiplied it by the average to get a new distance traveled column. Now each run had the same length and were thus comparable. Table 1 documents each segment and its average length.

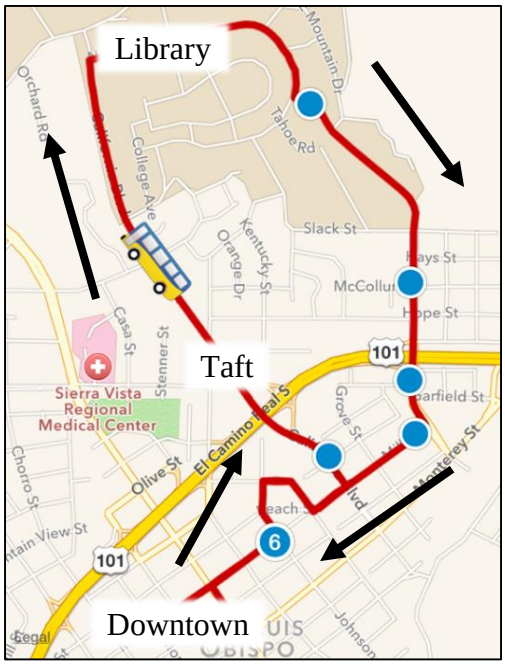


Figure 1: Route 6b

Table 1: Average Segment Length, in Miles

Route Segment	Description	Average Length	Cumulative Length
44	Library to Downtown	2.139	2.139
29	Downtown to Taft	0.842	2.981
0	Taft to Library	0.936	3.917

Now that I finally had usable timestamp and distance traveled data, I was able to create a preliminary time-space diagram in order to assess my results. This diagram pointed out to me several errors in my data that would have been hard to discover simply by looking at the spreadsheet of raw data. One such error was that the first run switched its segment number labeling. Instead of transitioning from Segment 0 to 44 to 29, it instead transitioned from Segment 44 to 0 to 29. Checking the coordinates revealed that the 0 and 44 were switched. Additionally, I found several instances where a segment label was misplaced. For instance, in a string of "44"s, there were an errant few "0"s. One other nuance that I discovered was that the system did not record duplicate locations. If a bus dwelled in the same spot for 10 minutes, then the timestamp would jump 10 minutes. However, my initial analysis had been treating each row as 15 seconds, and so this 10 minute jump would show up as only 15 seconds. I went back and added in rows to account for the missing intervals. With this completed, I concluded my error hunt and began plotting some charts, such as Figures 2 and 3 below.

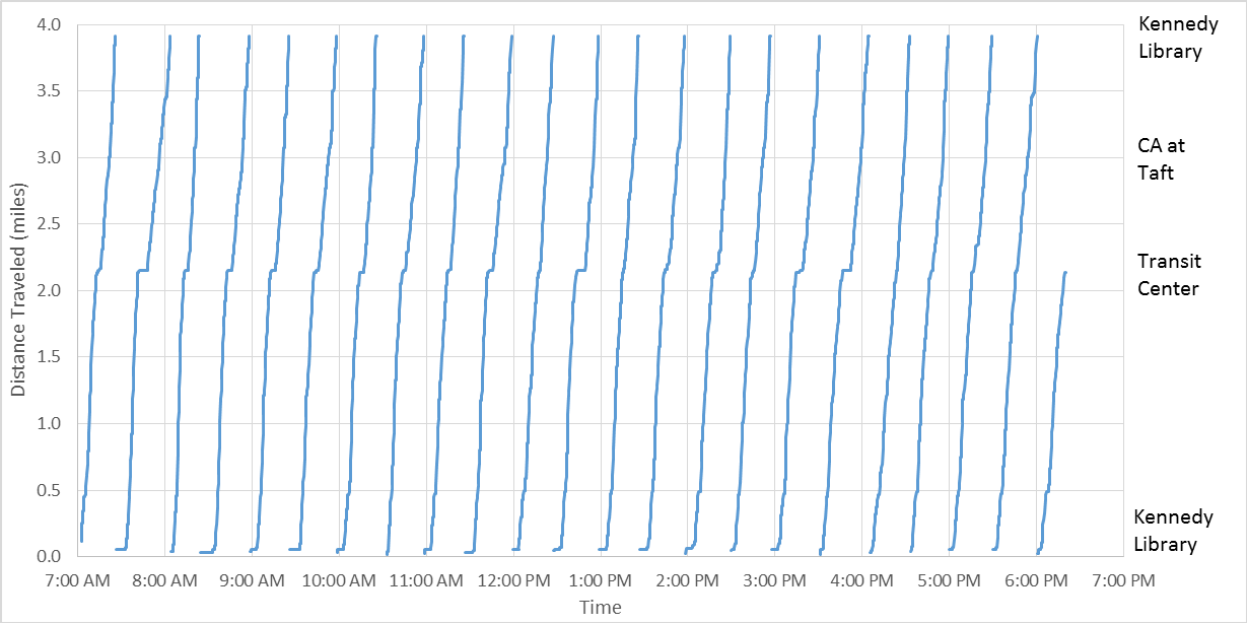


Figure 2: Time-Space Diagram Depicting Runs vs. Real Time

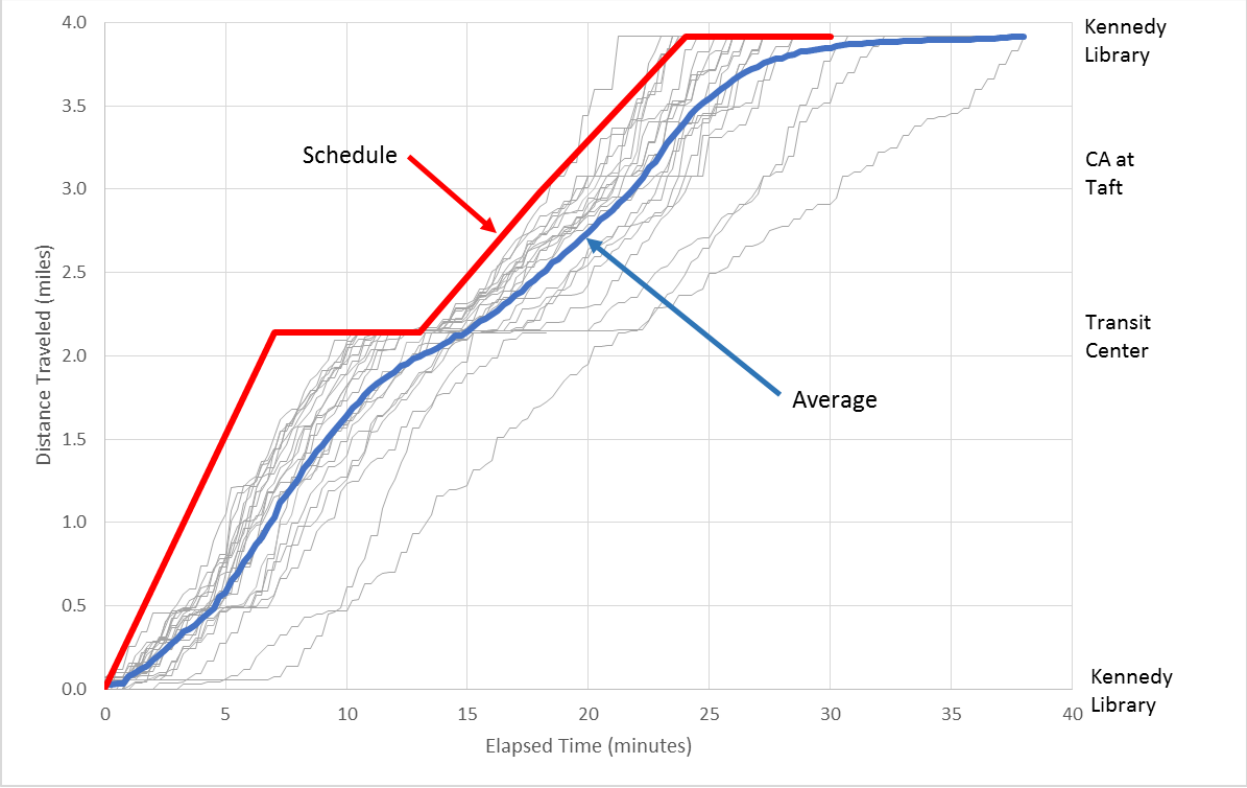


Figure 3: Time-Space Diagram Depicting Runs vs. Elapsed Time

I created Figure 2 first, but with 22 and a half runs spanning 11 hours, it became hard to see any appreciable trends other than varying dwell times at the Kennedy Library and Transit Center stops. To rectify this, I changed the X-axis to elapsed time per run, rather than clock time, with "zero minutes" correlated to the scheduled time that the bus is supposed to leave the library (either two or 32 minutes after the hour). With these runs compiled together, it was possible to add lines depicting how the runs compared versus the schedule and the average.

As one can see, pulling equations from this graph for anything other than the scheduled run is an exercise in futility. To get the equations, I had to go back to my data and compute them using the starting and ending point of each travel period. I found that the dwell equations were easy enough, as they are simply the distance of each stop along the route:

- Library (at start): $x = 0$
- Transit Center: $x = 2.139$
- Library (at end): $x = 3.917$
- x is distance, in miles

There was no dwell time for the California at Taft stop. In fact, since the bus rarely even stops at California at Taft, I merged the second two segments (29 and 0, depicting Downtown to Taft and Taft to the Library, respectively) into one segment for the latter part of my analysis. I then computed travel equations for each run in each direction, and computed the average:

- Average Library to Transit Center: $x = 11.2t - 0.750$
- Average Transit Center to Library: $x = 10.0t - 0.892$
- x is distance, in miles
- t is time, in hours

With the calculation of these equations, I completed my initial objectives. However, these equations are not very helpful and are hard to digest for the layperson. For this reason, I expanded my objectives to include finding dwell times and travel speeds, and comparing them to the schedule. Since I already had my calculations broken down by route segment to find the travel equations listed above, it was not too difficult to calculate the dwell times and travel speeds.

For the dwell times at the Kennedy Library stop, I found that my sample ranged from two to 11.25 minutes, with an average of 5.15 minutes. This is higher than the schedule, which has zero to six minutes, with an average of three minutes. For the schedule, since there is both an arrival and a departure time, and the bus can arrive any time between the two, it essentially allows a range, from "touch-and-go" where the bus arrives then immediately departs, to arriving right at the arrival time and sitting for all six minutes.

For the dwell times at the Downtown Transit Center stop, my sample returned 1.25 to 7.5 minutes of dwell time, with an average of 3.95 minutes. Like the Kennedy Library Stop, the Transit Center has a scheduled stop of zero to six minutes, with an average of three minutes. Therefore, the Transit Center is performing a lot better regarding keeping the dwell time within schedule, with most runs stopping for less than six minutes.

For the travel speed from the Library to the Transit Center, the sample ranged from 7.5 to 16.7 mph, with an average of 11.2 mph. This is slower than the scheduled 9.9 to 18.3 mph, with an average of 14.1 mph. The schedule has a range for a similar reason as the dwell time: the lower speed is if the bus takes all six minutes of scheduled stop time to arrive at the stop, and the higher speed is if it arrives right on the listed arrival time.

For the travel speed from the Transit Center to the Library, the runs were slightly faster than the schedule, with a range of 6.2 to 13.2 mph versus the schedule's 6.7 to 9.7 mph. Interestingly, the sample's average speed was noticeably higher than the schedule: 10.0 mph versus 8.2 mph.

CONCLUSIONS AND RECOMMENDATIONS

In conclusion, the bus dwelled longer than scheduled at the Library, and less than scheduled Downtown. Therefore, the schedule should have less dwell time Downtown and more at the library. Additionally, the bus traveled slower than scheduled heading from the Library to Downtown, and faster than scheduled from Downtown back to the Library. For this reason, the schedule should have less travel time from Downtown to the Library, and more from the Library to Downtown.

To help the schedule better reflect reality, I recommend the following changes, outlined in Table 2:

Table 2: Proposed Dwell and Travel Time Changes (in Minutes)

	Library Dwell	Library to Downtown	Downtown Dwell	Downtown to Library
Old	6	7	6	11
New	7	9	5	9
Difference	+1	+2	-1	-2

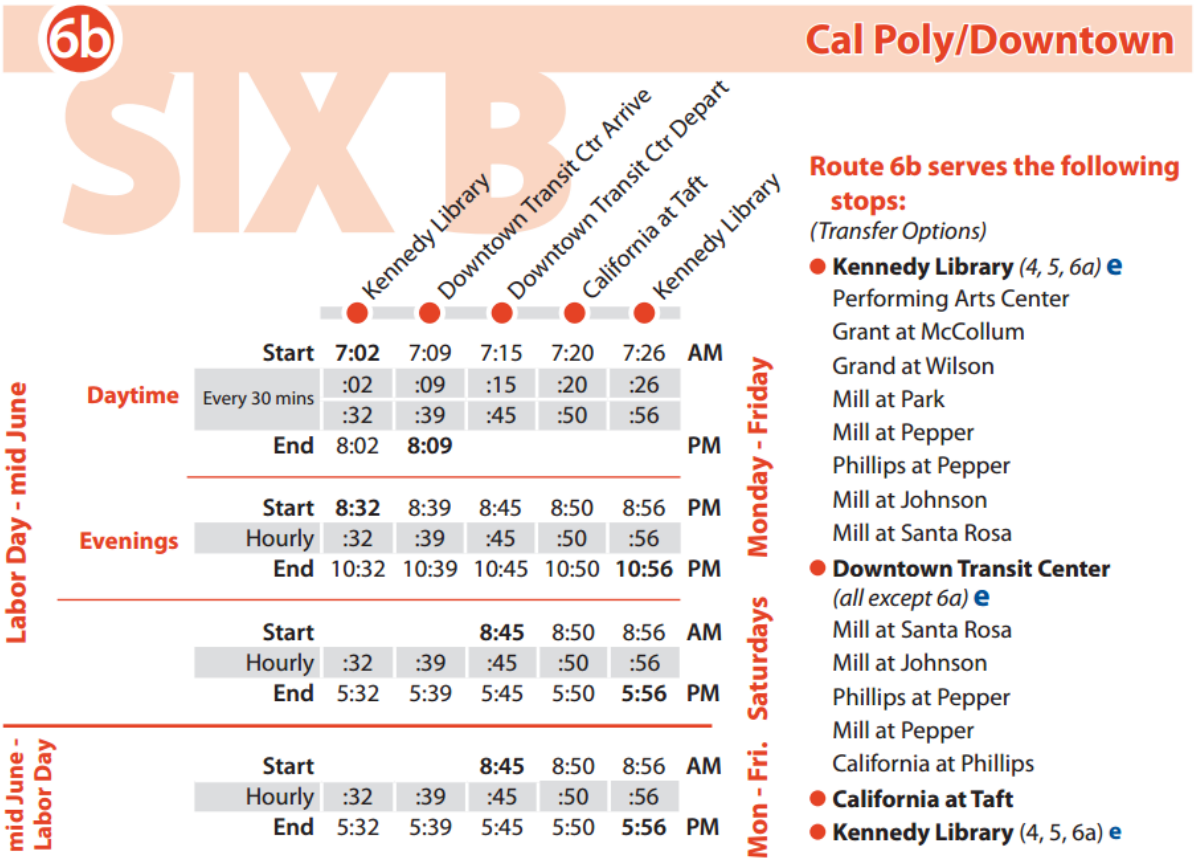
These changes manifest themselves in the schedule, as Table 3 depicts:

Table 3: Proposed Schedule Changes

	Kennedy Library Depart	Transit Center Arrive	Transit Center Depart	California at Taft	Kennedy Library Arrive
Old	:02	:09	:15	:20	:26
	:32	:39	:45	:50	:56
New	:02	:11	:16	:20	:25
	:32	:41	:46	:50	:55
Difference	0	+2	+1	0	-1

These changes to the schedule will help ensure that the bus runs more reliably. It is easier to match the schedule to the bus, rather than the bus to the schedule. The former requires only minor schedule adjustments that carry through to posted signage and similar channels, whereas the latter typically requires more expensive solutions, such as transit signal priority implementation, stop consolidation, or hardscape improvement. With this simple desk analysis, a transportation analyst can assess the reliability of a bus route and suggest minor tweaks to improve service to and goodwill among the riding public.

APPENDIX A: ROUTE 6b SCHEDULE



6b

Cal Poly/Downtown

SIX B

Kennedy Library | Downtown Transit Ctr Arrive | Downtown Transit Ctr Depart | California at Taft | Kennedy Library

Labor Day - mid June

Monday - Friday

Daytime	**Start**	**7:02**	7:09	7:15	7:20	7:26	AM
	Every 30 mins	:02	:09	:15	:20	:26	
		:32	:39	:45	:50	:56	
	End	8:02	**8:09**				PM
Evenings	**Start**	**8:32**	8:39	8:45	8:50	8:56	PM
	Hourly	:32	:39	:45	:50	:56	
	End	10:32	10:39	10:45	10:50	**10:56**	PM

Saturdays

Start			**8:45**	8:50	8:56	AM
Hourly	:32	:39	:45	:50	:56	
End	5:32	5:39	5:45	5:50	**5:56**	PM

mid June - Labor Day

Mon - Fri.

Start			**8:45**	8:50	8:56	AM
Hourly	:32	:39	:45	:50	:56	
End	5:32	5:39	5:45	5:50	**5:56**	PM

Route 6b serves the following stops:

(Transfer Options)

- **Kennedy Library** *(4, 5, 6a)* e
  - Performing Arts Center
  - Grant at McCollum
  - Grand at Wilson
  - Mill at Park
  - Mill at Pepper
  - Phillips at Pepper
  - Mill at Johnson
  - Mill at Santa Rosa
- **Downtown Transit Center** *(all except 6a)* e
  - Mill at Santa Rosa
  - Mill at Johnson
  - Phillips at Pepper
  - Mill at Pepper
  - California at Phillips
- **California at Taft**
- **Kennedy Library** (4, 5, 6a) e

Source: SLO Transit