PUBLIC TRANSIT RIDERSHIP

ANALYSIS OF THE UNITRANS TRANSIT SYSTEM FOR RESOURCE OPTIMIZATION

Table of Contents

Executive Summary

The UC Davis Unitrans transit system is facing a convergence of several problems that can dramatically affect its operations. Unitrans continues to see excessive ridership during inclement weather leading to increasingly dissatisfied customers. Second, a legislatively mandated minimum wage increase in California will effectively increase their labor cost by $800,000 annually starting in 2021. Finally, the Unitrans fleet is changing with the addition of three double decker buses that need to be scheduled effectively. Unitrans needs a plan to address these issues and has asked HAT Consulting to make an analytical evaluation and make recommendations for operational change.

A statistical analysis of Unitrans operations performed by HAT Consulting has developed forecast models that inform Unitrans of key factors that influence operational decisions. These key insights can inform Unitrans of peak demand periods throughout the year, and passenger demand variations based on prevailing weather patterns. Unitrans can utilize this information for various operational changes such as decreasing the service levels during periods of low demand and result in a savings of over $33,000 in operational costs. Alternatively, Unitrans can increase hiring by 6% to accommodate peak demand periods along with doubling bus line capacity by shifting bus resources. With HAT Consulting recommendations, Unitrans can fully optimize its resources and budgetary decisions.

1. Introduction

Founded in 1968, with two vintage double decker buses from London, Unitrans is the public bus system for the University of California Davis (UC Davis) and the City of Davis. With 48 buses and 18 routes, Unitrans carries over 4 million passengers per year[1]. Over 22,000 passengers use the bus system on a normal day. The drivers, supervisors, and much of the support staff for Unitrans are UC Davis students providing transportation to students, and community members as they travel to downtown Davis, schools, hospitals, shopping centers, theatres and many other destinations.

Unitrans receives its revenue from various sources. The bulk of the revenue comes from the Associated Students of UC Davis in the form of a Transit Fee. For fiscal year 2016-2017, the fee provided $2,574,746. Other sources included $710,000 from the City of Davis, $20,000 from Yolo County, $1,300,000 from Federal funding, $265,000 from Fares (estimated), $31,000 from advertising and $170,000 from miscellaneous sources[2]. Since funding from ridership fares account for such a small part of Unitrans' revenues, it does not need to heavily rely on it to fund its operations.

Despite having fairly stable sources of revenue, Unitrans is facing a number of issues. The transit system is facing a growing annual deficit coming from their operational labor costs. Legislation that incrementally raises the California minimum wage to $15 per hour by 2021 is the biggest driver of this deficit. Unitrans estimates a $200,000[3] annual increase in labor costs through 2021. Additionally, Unitrans continues to experience over capacity ridership during peak demand periods coinciding with inclement weather. This has led to crowded buses and unhappy passengers. Unitrans uses two "Tripper[4]" buses to help alleviate the crowding, but it is still insufficient. Lastly, Unitrans plans to replace

[1] http://unitrans.ucdavis.edu/about/

[2] Palmere, A. pp. 5-6.

[3] Palmere, A. pp. 6.

[4] A "Tripper" bus is a spare bus in the Unitrans fleet that is deployed to crowded bus lines as needed to increase capacity. Unitrans currently keep 2 single deck buses in reserve to fill this role.

three of their regular buses with double decker buses with a capacity of 100 passengers each. Unitrans needs a way to efficiently utilize their resources and plan for future operations to increase customer satisfaction.

HAT Consulting volunteered to analyze Unitrans' ridership data to determine predictive forecast models and prescribe recommendations to optimize the transit system's operations. HAT Consulting concentrated on determining peak ridership periods and variations in ridership due to external factors. From the analysis, HAT Consulting has determined that throughout the year, Unitrans experiences several periods of peak ridership of over 26% that are caused by conditions such rain resulting in more passengers to ride the bus or periods of the academic year such as the beginning and end of quarters that also increase ridership. During these weeks, Unitrans should shift their double decker buses to high use lines, effectively increasing capacity from 60 passengers up to 200 passengers per run. HAT Consulting has found that Unitrans faces three climate scenarios: El Nino, normal years, and drought years which affects their overall ridership. With this, Unitrans can plan to increase their budgets by 6% for operational labor in support of increase passengers in high demand years or they can reduce service levels by 2.5% in low demand years resulting in savings of over $33,000 and reducing wear and tear on the bus fleet.

The remainder of this report describes the methods of analysis that HAT Consulting performed for Unitrans and is broken into 4 sections; first, is the examination of the raw data provided by Unitrans to observe possible trends and any factors that may affect the transportation system's ridership. Second, a forecast model using decomposition methods is determined and tested. Third, regression techniques are used to create a forecast model for Unitrans ridership. Finally, recommendations and action steps for Unitrans based on the findings and predictions of the models are provided to Unitrans by HAT Consulting.

2. Data Characteristics

In this section, the data used for the analysis will be discussed. Information on the data will be stated for each variable considered, the reduction of the data, the compiling of the data, data observations and correlations.

2.1 Data Background

Figure 1: Unitrans Route Map

HAT Consulting received raw data from Unitrans from the time period between January, 2014 and May, 2017. A total of forty-five Excel files were received and each included data on: date of service, time of service, bus identification number, bus stop location, bus stop identification, trip route and identification number, and the number of boarding and de-boarding passengers for each stop location. This accounted for a total of three million data points to be evaluated. To illustrate how so much data is collected consider Figure 1 which shows the bus routes for weekday services, and the oval sections which highlight the areas that will be the focus of the analysis. The highlighted region contains the G-, J-, and W-line trip routes which were the top three routes in regard to the total number of passengers. These routes have the largest ridership because along their path they have a high number of rental apartments where most of the tenants are students. A note on the data provided; there were periods of zero ridership during school breaks such as spring break, and holiday break. The zero ridership varies year-to-year, and is mainly attributed to the Unitrans management who decide if buses will run during those times. To compensate for the zero ridership, historical averages replaced those data points.

2.2 Weather and Precipitation

In addition to bus routes data, the daily temperature and precipitation was added to the model. Weather data for the years evaluated were collected from the California Irrigation Management Information System (CIMIS)[5].

2.3 Data Reduction

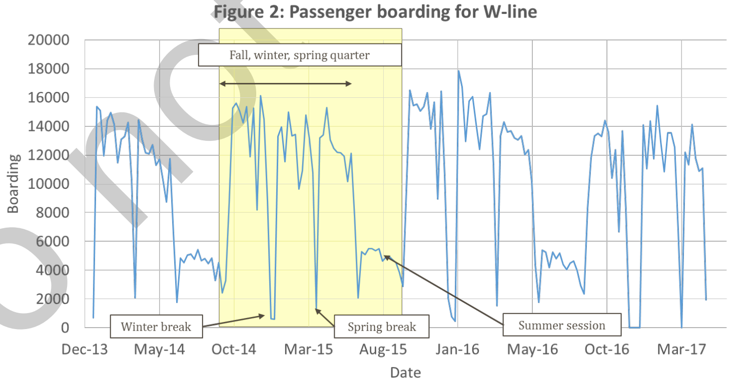
In reducing the bus lines, only the date of service and the number of passenger boarding was considered. Since we only focus on total ridership, information regarding specific stops, bus used and de-boarding (which directly correlates to boarding's) were unnecessary to our evaluation. The passengers were then clustered, and summed, into groups of seven days to capture ridership by week. Each week was uniquely identified with an index to account for all fifty-two weeks in a year. Similarly, the weather data was grouped into weekly clusters. The daily temperature data was averaged for each week, and the weekly rain was summed up. This data was used for all three bus lines in the analysis.

2.4 Compiling Data

With the data reduction complete, three data sets were created for each of the bus lines in consideration. In summary, the final data was formatted the same and included: starting week of service number of passenger boarding for each bus line (G-, J-, or W-line), average temperature, weekly total rainfall, week numerical value to track week of observation, and weekly indicator.

2.5 Observations of Bus Lines

Figure 2: Passenger boarding for W-line



With our focus on three bus lines, we begin our efforts to observe any trends for the timeframe we are evaluating. Figure 2 is an example of annual passenger count for the W-line, and similar plots for the G- and J-line can be reviewed in Appendix A-1. The figure highlights the 2015 school year, identifying major dates such as duration of the quarter session, winter break, spring break,

[5] CIMIS is a database that is integrated into the University of California Statewide Integrated Pest Management (UC IPM) program. CIMIS was developed by the California Department of Water Resources and the University of California, at Davis. It was designed to assist irrigators in managing their water resources more efficiently.

and summer session. The troughs are indicators of holidays that are recognized by the university, like Veteran's Day, Thanksgiving Day, Martin Luther King Day, to name a few. The dates that are most impactful to the ridership are winter break and spring break. This observation is attributed to the length of the break being at least one week, so students are more likely to be out of town. These seasonal trends are also observed for the portion of school year 2014, school year 2016, and part of 2017 to-date.

2.6 Data Correlation

Temperature and rainfall was included in our evaluation to determine how weather impacted ridership, and that is shown in Figure 3. The temperature and rainfall are seasonal, and to help better understand trends, the correlation values were calculated, and provided in Table 1 for boarding with temperature, and boarding with rainfall. This was completed for all three bus lines and each had similar values. The correlation values suggested that boarding and temperature had an inverse relationship, so when the temperature drops and it's cold, ridership increases and vice versa. As for passenger and rainfall, it had a direct relationship meaning that as rainfall increases so does ridership, and when rainfall drops, so does ridership.

Figure 3: Passenger Boarding, Temperature, and Rain for W-line

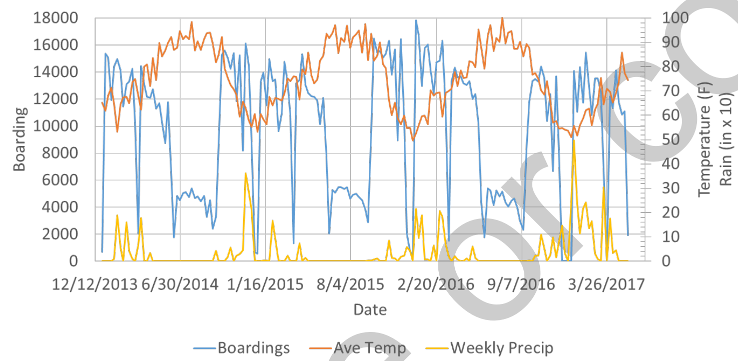


Table 1: Correlation Values for each line.			
	G-Line Boarding	J-Line Boarding	W-Line Boarding
Temp	-35.4%	-34.6%	-32.7%
Rain	26.8%	26.8%	22.5%

The direct relationship for rainfall and passengers makes sense since good weather allows for alternate methods for students to commute to campus, like riding their bicycles or even walking.

In the next section, we will discuss the model our model selection process, the model that best fits our data, and provide internal and future forecasts.

3. Model Analysis

After reviewing the data and its characteristics we moved into choosing the appropriate model for the purpose of forecasting ridership for the next 52 weeks. This forecast will be used for recommendations on bus services and staffing levels. Due to the high levels of seasonality in ridership, great care was taken in selecting the model which both fit the data best and its use could be replicated across all lines of services for consistent forecasting.

3.1 Model Selection

Multiple models were considered for the best potential forecast. The Mean Absolute Error[6] (MAPE) was used to narrow down the selection (see Appendix B-1). Other than Winters model and Multiple Regression, all models produced unacceptable errors. Due to close similarities in error values the Winters model and Multiple Regression models were chosen for expanded evaluation. To make a

[6] The Mean Absolute Percentage Error is a measure of the average of absolute distance between errors and actual or predicted values. This ratio allows for comparison of models. Lower values indicate a more accurate model. Delurgio, 1998. pp. 55-56.

determination of which model to use we conducted an internal forecast to measure which model performed best.

We performed the same internal forecasts for the Winters model and Multiple Regression models on all lines of service we were evaluating.

3.1.1 Winters Model

First, we did an internal forecast on all three bus lines using 48 weeks of data. For the Winters model, only the total number of weekly passengers over time were utilized in building the model and the forecast. We created a forecast for bus lines J (Figure 4). Line J shows a clear lack of fit for the forecast and the subsequent error calculations confirm this finding. We performed the same analysis for lines G and W (Appendix B). Line G is similar to Line J although the fit is slightly better. The errors associated with Line G were higher than we would have preferred. Lastly, we performed the forecast on Line W. For this line our fit was significantly better. The error values were also significantly better for this bus line. Despite the issues with the first two forecasts we thought it was possible that the Winters modeling method would be useful for our purposes.

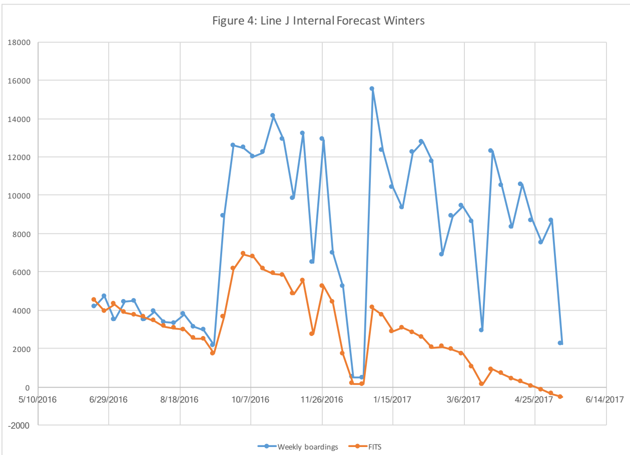


Figure 4: Line J Internal Forecast Winters

3.1.2 Multiple Regression Model

Again, just as with the Winters model, we performed forecasts on 48 weeks of ridership data for Lines J, G and W. The multiple regression differs from the Winters model in that there are more independent variables considered in the creation of the model. In addition to the number of passengers per week, we utilized average weekly temperature (°F), and total rain fall (inches), all over time. Since we were simulating a forecast, we used averages for both temperature and rainfall instead of actuals for each week. This methodology aligns with how we would perform the actual forecast and therefore give us the best idea of how well the model functions. We started with Line J and it was quickly apparent that it fit the forecast significantly better (Figure 5). The error calculations were also low. We performed the same analysis for lines G and W (Appendix B-2 and B-4). We evaluated Line G and found a similar result to line J. The predicted values were lower than the actuals, however, they follow the weekly trend nicely and the calculated errors were not a cause for concern. Lastly, we evaluated Line W and saw a similar trend to Line G except instead of under estimating slightly the model is over estimating slightly.

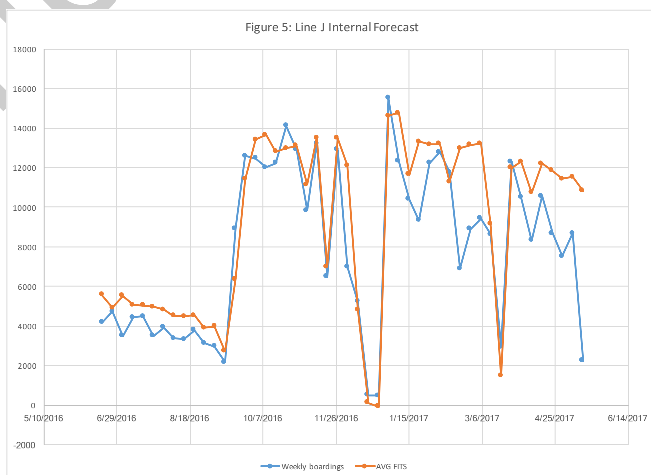


Figure 5: Line J Internal Forecast

Now that we have evaluated both models we will determine the best model for our purposes.

3.1.3 Model Comparison

As mentioned previously, all three bus lines exhibited similar errors during the initial model building for both Winters model and Multiple Regression. We therefore decided that the model which performed best at the internal forecast would be the best fit for our analysis. Error comparisons can be found in Appendix B-4. One of the main differences between the Winters model and the Multiple Regression is that the Multiple Regression consistently followed the weekly trend where the Winters model was unreliable for two of the three bus lines. Despite Lines G and W having a slightly better MAPE value for the Winters model, we ultimately decided to utilize the Multiple Regression model for the following reasons:

- High correlation between ridership and temperature and rain.
- Consistent internal forecasts for all three lines.
- More easily replicable across all Unitrans lines.

Now that we had chosen a model we needed to build the full models and interpret the results.

3.2 Model Interpretation

After selecting the Multiple Regression model for our final forecast we used all the data points to create a full, robust formula for prediction. The truncated (full equations can be found in Appendix B-3) equations are as follows:

- Line J Model

  Weekly Boardings = 1,410 - 11.1*Ave Temp - 5.65*Time + 874*Weekly Precipitation + Weekly Index*Week.

- Line G Model

  Weekly Boardings = 3,473 - 45.8*Ave Temp - 10.88*Time + 342*Weekly Precipitation + Weekly Index*Week.

- Line W Model

  Weekly Boardings = -1 + 1.8*Ave Temp + 8.82*Time + 722*Weekly Precipitation + Weekly Index*Week.

The wide ranges of ridership on a week to week basis, which closely correlates to the UC Davis academic calendar, caused us to evaluate the model on a week by week seasonality basis. The seasonality values were consistent in scale for each equation. For example, we observed higher values in weeks 5, 6 and 7 across all models, which corresponds to the beginning of the spring quarter. Both Lines J and G are negatively impacted by an increase in temperature where Line W is basically neutral. All three Lines experience an increase in ridership during weeks of heavy rain. Lastly, lines J and G are experiencing a decline in total ridership over time as indicated by the negative variable for Time. Conversely Line W is experiencing an increase in ridership over time.

Next we wanted to ensure that our data fit our assumptions of linearity, homoscedasticity and other diagnostic measures.

3.3 Model Diagnostics

To ensure there were no surprises with the data we performed additional data diagnostics (See Appendix A-2). The data tested within all appropriate ranges. We therefore determined that it was appropriate to move forward with our analysis.

Next, we will use these models to determine the expected ridership by week for the three bus lines.

3.4 Forecast

Now that we have established our prediction models we can accurately determine the ridership over the next 52 weeks (week 5/21/2017 to week 5/13/2018). Since we do not have actual temperature and precipitation data for the future 52 weeks we developed models for these values which look at the average, maximum (high temperature/low rainfall) and minimum (low temperature/high rainfall) values for each variable. We have labeled the minimum as El Nino years and maximum as drought years. Intuitively, a higher average temperature in Davis, CA would correspond to lower precipitation values and vice versa. Using the equation for Line W we forecast 52 weeks ahead (Figure 6). It can be observed that during the summer where temperatures are high and school is out of session that all three forecasts follow each other closely. The lowest point of ridership comes during Christmas break when school activity is at its lowest and many students have gone home for the holidays.

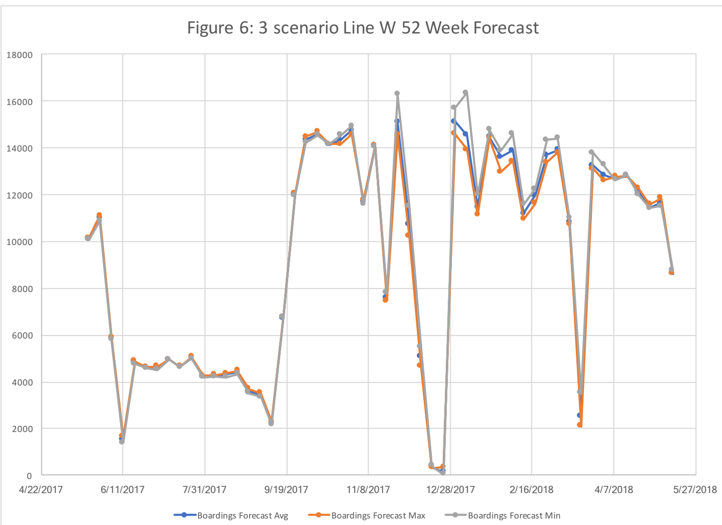


Figure 6: 3 scenario Line W 52 Week Forecast

Using this information we are able to make the appropriate recommendations for Unitrans in how to best maximize their operations.

4. Summary and Recommendations

From our analysis, we have three climate based scenarios for Unitrans to consider. We categorize these as: El Nino (wet) year, Normal year, and Drought year. An El Nino year corresponds to our situation where there is low temperatures and high rain fall throughout the year. A normal year contains average temperatures and rainfall. A drought year is where there are high temperatures and low rainfall. These climate conditions directly correspond to the three forecasts scenarios that we explored. Unitrans can expect to experience El Nino years every 2-7 years[7], or 4 years on average. Drought years occur on a similar cycle and are characterized by the effect called "La Nina[8]" that pushes precipitation north causing a dryer season in California. These three climate situations can be readily gotten from long term meteorological predictions for each year and Unitrans can use the information to adjust their operations as necessary.

---

[7] https://www.wildlife.ca.gov/conservation/marine/el-nino
[8] https://www.nationalgeographic.org/encyclopedia/la-nina/

Our Analysis and forecasts also reveals specific weeks where Unitrans will experience higher than average ridership on the J, G and W lines. Between each of the three forecast scenarios across all three lines, we have found key weeks where ridership was 26% or higher than average. These peak weeks not only correspond to periods of inclement weather, but also to unique periods of the school year. For example, weeks 2, 15, and 41 represent the beginning of a quarter when students return and ridership increases. A similar increase occurs towards the end of a quarter when exams are scheduled in weeks 11 and 49. A full listing of these peak weeks can be found in appendix B-7.

Now that we have established some key insight into the potential future operation of Unitrans' top three lines, we have two types of augmentations that we recommend. The first augmentation is regarding Unitrans' annual budget. Because of the increase in minimum wage, Unitrans must increase their budget by $200,000 each year to cover their operational labor. We recommend increasing this amount by 6%, or $212,000, each year until 2021. The 6% increase corresponds with the difference in ridership from the lowest scenario to the highest ridership scenario of a wet year. This will allow Unitrans to increase their driver pool to cover the increase in demand. However, we do not recommend Unitrans increase their workforce immediately. The 6% increase should be placed in reserve and when an El Nino year is predicted, Unitrans should actively recruit and train additional drivers to fill the extra need. Once the El Nino year passes, Unitrans can allow their work force to reduce to normal levels from attrition as drivers graduate and leave the university.

Because of our recommendation, Unitrans will have a budget gap of $848,000 starting in 2021. To close this gap, we have a number of options that Unitrans should take. First is an incremental increase in fares from $1 to $2 that should be complete by 2021. The $2 fare will generate an additional $265,000 in revenue assuming current paid ridership stays the same. The increased fare is a competitive fare as other local transit systems such as Sacramento RT or San Francisco MUNI have fares of $2.75. Unitrans would still be a low cost option. This leaves $583,000 that need to found. Unitrans can lobby the ASUCD to cover this gap by raising the Transportation Fee by $16 per student. This is a fairly reasonable request, but if Unitrans would like to be sensitive to every increasing student fees, they can seek to find additional funding from other means. This can include raising their advertising fees to generate higher ad revenue. Unitrans can also apply for higher funding form the Federal transit program. Additionally, they can lobby the City of Davis and Yolo County to increase their contributions as well. In general, Unitrans has many options to cover their increase in operational labor costs.

The second augmentation that we recommend to Unitrans is regarding their bus schedules and operations. These recommendations derive from the three climate scenarios that they can face.

El Nino Year

- Switch to a 3 or 4 tripper bus system. This will allow for more flexibility to increase capacity on demand.
- Move double decker buses from other lines to J and G line during peak periods. This will increase capacity on a run from 120 passengers to 160 or 200 passengers depending on bus combinations.
- Run two tripper buses on peak demand for the W-line.
- Increase maintenance cycle to allow the extra trippers to be available during these peak periods.

Normal Year

- Same operational changes as El Nino Year.

- However, depending on observed demand, Unitrans may not need to move double decker buses to the J and G line and use single deck trippers instead. This would allow capacity on other lines that use the double decker buses to not to be diminished.

Drought Year

- Consider reducing the number of bus runs on J, G and W line by 2.5%. This corresponds to the reduced forecasted ridership in a drought year compared to a normal year.
- This reduction will achieve a $33,296 savings in operation costs.

Due to the limited scope of our report and analysis, we have some next steps for Unitrans to take. First, Unitrans should perform a similar analysis as ours on the other bus lines in their system. We observed a reduction in ridership across the J, G and W lines but an increase in ridership across the entire transit systems. Further analysis will determine which bus lines are contributing to this increase and modeling based on those bus lines will give Unitrans more tools to better utilize their resources. Additionally, Unitrans should re-evaluate our models on an annual basis to integrate new ridership and weather data, further improving the accuracy of the forecasts from the models.

By following our recommendations and next steps, we feel that Unitrans can greatly improve their resource utilization. With a more efficient transit system, Unitrans will have more customer satisfaction and increased ridership.

Appendix

This appendix contains additional information regarding data characteristics and supporting information for our model and forecast predictions.

Appendix A- Supplemental information for Model Structure and Diagnostics

A-1 Additional Trend Data for Bus G- and W-Line

Figure A-1 and Figure A-2 are provided to show that the trend observed for ridership of the G-line and J-line, respectively. Both the G- and J-line have similar trends to the W-line which was presented in the data characteristics section.

Figure A-1: Passenger boardings for the G-Line.

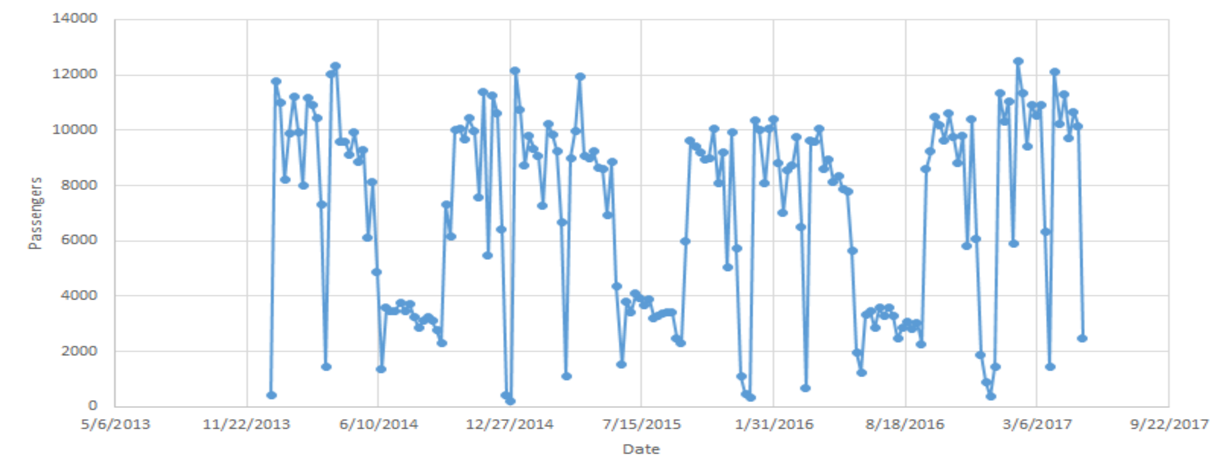
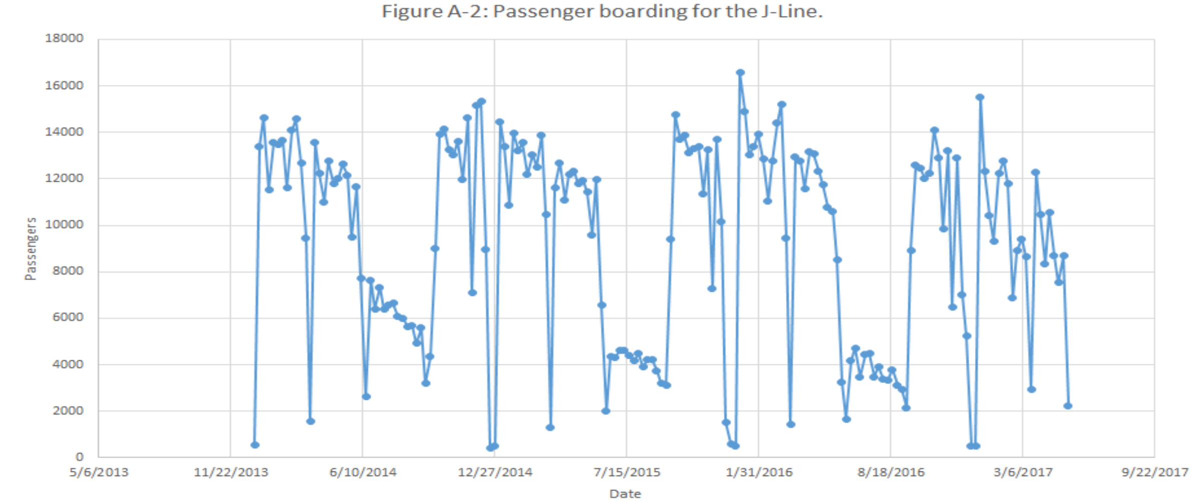


Figure A-2: Passenger boarding for the J-Line.



A-2 Assumption Testing of Data Set for Regression

In order for a linear regression model to be valid, there are several assumptions about the data that we must take. In this section, we test the following assumptions:

1. Normality
2. Homoscedasticity (Constant Variance)
3. Linearity
4. Independence
5. Multicollinearity

We found that all assumptions have been met without amendments to the data. A summary is presented in the following table:

Table A-1 Data assumptions for linear Regression Modeling		
Assumption	**Test Status**	**Comment/test used**
Normality	Pass	Observation of bell shaped histogram of residuals[9]
Homoscedasticity	Pass	Residual Plot evaluation
Linearity	Pass	Scatter Plot of imports vs time evaluation
Independence	Pass	Durbin & Watson Test[10]

Appendix B – Supporting information for Model and Forecast

In appendix B, we present the models that were considered for the forecast of the Bus ridership as well as the model that was chosen, Multiple Regression Model. Additionally, we present a comparison of the models to show why we chose our model. In appendix B-5, we present the data used to forecast bus ridership as well as the average forecasted values for May 2017 through May 2018.

B-1 Model Comparisons

Several models were considered during the course of HAT consulting's analysis of Unitrans data sets. The Mean Absolute Percentage Error[11] (MAPE) was used to compare the relative accuracy of each forecasting model. By examining the ratio of error produced by each model compared to forecasted values, a more accurate determination of model validity can be attained. HAT Consulting determined that the Winter's model and Multiple Regression models were the strongest candidates for further evaluation. The ARIMA model was not a viable solution due to a significant decrease in ridership during spring break every spring semester which prevented the ARIMA model to not demonstrate any significance.

Table B-1 Model MAPE Comparison	
Model	**MAPE**
Simple Regression	137%
Moving Average	58%
Holt's Method	124%
Winters Method	11%
Multiple Regression	11%
ARIMA	N/A

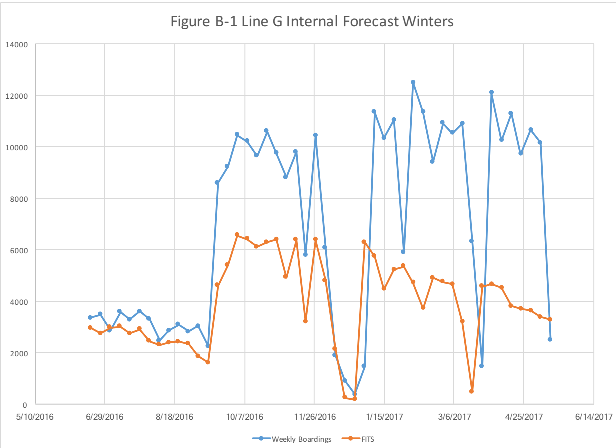
B-2 Winter's Model Internal Forecasts for Lines G and W

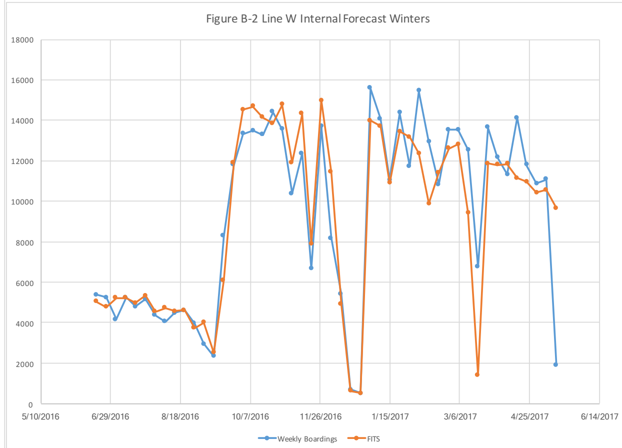
An internal forecast was performed using the Winter's model for the G and W line data sets to determine the accuracy of the forecast model. HAT Consulting withheld 48 weeks of observed data and forecasted values were compared to those observations for accuracy.

[9] Residuals are the difference between actual values of y and the values calculated by the regression line. Keller, 2012. Pp. 650.

[10] "The Durbin-Watson test allows the statistics practitioner to determine whether there is evidence of first-order autocorrelation." Keller, 2012. Pp. 716-719

[11] The Mean Absolute Percentage Error is a measure of the average of absolute distance between errors and actual or predicted values. This ratio allows for comparison of models. Lower values indicate a more accurate model. Delurgio, 1998. Pp. 55-56.





B-3 Multiple Regression Model

Line J Full Model Equation

Weekly Ridership = 1410 - 5.65 Time - 11.1 Ave Temp + 874 Weekly Precip + 13717 Week 2 + 13599 Week 3 + 10868 Week 4 + 13219 Week 5 + 12036 Week 6 + 12582 Week 7 + 11095 Week 8 + 12514 Week 9 + 12797 Week 10 + 13015 Week 11 + 9350 Week 12 + 917 Week 13 + 11924 Week 14 + 12075 Week 15 + 11038 Week 16 + 12387 Week 17 + 12299 Week 18 + 11837 Week 19 + 12016 Week 20 + 11170 Week 21 + 9777 Week 22 + 10733 Week 23 + 5874 Week 24 + 2068 Week 25 + 5896 Week 26 + 5310 Week 27 + 5838 Week 28 + 5429 Week 29 + 5395 Week 30 + 5404 Week 31 + 5151 Week 32 + 4893 Week 33 + 4875 Week 34 + 4921 Week 35 + 4266 Week 36 + 4405 Week 37 + 3053 Week 38 + 6611 Week 39 + 11728 Week 40 + 13711 Week 41 + 13909 Week 42 + 12960 Week 43 + 12746 Week 44 + 12894 Week 45 + 11216 Week 46 + 13460 Week 47 + 6679 Week 48 + 12440 Week 49 + 11065 Week 50 + 3689 Week 51 - 18 Week 52.

Line G Full Model Equation

Weekly Ridership = 3473 - 45.8 Ave Temp + 342 Weekly Precip - 10.88 Time + 10968 Week 2 + 10388 Week 3 + 8125 Week 4 + 9975 Week 5 + 9727 Week 6 + 9355 Week 7 + 7625 Week 8 + 10100 Week 9 + 9854 Week 10 + 9961 Week 11 + 7319 Week 12 + 1487 Week 13 + 10451 Week 14 + 11209 Week 15 + 11326 Week 16 + 9774 Week 17 + 10043 Week 18 + 9832 Week 19 + 9675 Week 20 + 9528 Week 21 + 8060 Week 22 + 8995 Week 23 + 5282 Week 24 + 2681 Week 25 + 4949 Week 26 + 4949 Week 27 + 4944 Week 28 + 5206 Week 29 + 4866 Week 30 + 5425 Week 31 + 4391 Week 32 + 4431 Week 33 + 4550 Week 34 + 4779 Week 35 + 4570 Week 36 + 4229 Week 37 + 3468 Week 38 + 7830 Week 39 + 8960 Week 40 + 11001 Week 41 + 10781 Week 42 + 10141 Week 43 + 10387 Week 44 + 10300 Week 45 + 8022 Week 46 + 10402 Week 47 + 5357 Week 48+ 10011 Week 49 + 7689 Week 50 + 3179 Week 51 + 363 Week 52.

Line W Full Model Equation

Weekly Ridership = -1 + 1.8 Ave Temp + 722 Weekly Precip + 8.82 Time + 14407 Week 2 + 14422 Week 3 + 11087 Week 4 + 14390 Week 5 + 13287 Week 6 + 12913 Week 7 + 10359 Week 8 + 11870 Week 9 + 13048 Week 10 + 13392 Week 11 + 9915 Week 12 + 725 Week 13 + 12519 Week 14 + 12650 Week 15 + 12941 Week 16 + 12083 Week 17

+ 12029 Week 18 + 11394 Week 19 + 11628 Week 20 + 10465 Week 21 + 9587 Week 22 + 10515 Week 23 + 5364 Week 24 + 1036 Week 25 + 4443 Week 26 + 4170 Week 27 + 4641 Week 28 + 4645 Week 29 + 4410 Week 30 + 4778 Week 31 + 3986 Week 32 + 4173 Week 33 + 4026 Week 34 + 4084 Week 35 + 3198 Week 36 + 3453 Week 37 + 1927 Week 38 + 5562 Week 39 + 11841 Week 40 + 14558 Week 41 + 14837 Week 42 + 14255 Week 43 + 13844 Week 44 + 14803 Week 45 + 11957 Week 46 + 14569 Week 47 + 7640 Week 48 + 14155 Week 49 + 10613 Week 50 + 3653 Week 51 - 256 Week 52.

B-4 Comparison of Winter's Method and Multiple Regression

Error values were used as a means to compare the Winter's model and Multiple Regression Model. These error help to determine which of the two models may provide the most accurate forecasts. Errors used for comparison was the Mean Error (ME) which helps to determine if the models are under or over forecasting; Mean Square Error (MSE) is the average of sum of squared errors; Mean Absolute Deviation (MAD) is the is a measure of error dispersion that is less sensitive to outliers. MSE and MAD take with MAPE, provides a clearer picture of forecast accuracy[12].

ME Values	Multiple Regression Internal Forecast	Winter's Method Internal Forecast	Multiple Regression External Forecast (Drought Scenario)	Winter's Method External Forecast
J-Line	0.0000	8.78895	-1299.15	4978.19
G-Line	0.0000	14.9356	-753.852	3014.45
W-line	0.0000	-1.3772	-1441.39	61.1014

MSE Values	Multiple Regression Internal Forecast	Winter's Method Internal Forecast	Multiple Regression External Forecast (Drought Scenario)	Winter's Method External Forecast
J-Line	880,315	1,143,903	5,707,547	38,733,127
G-Line	419,123	680,349	7,766,942	18,201,525
W-line	962,107	1,706,536	5,883,683	3,573,154

MAD Values	Multiple Regression Internal Forecast	Winter's Method Internal Forecast	Multiple Regression External Forecast (Drought Scenario)	Winter's Method External Forecast
J-Line	678.737	713	1691.28	5033.53
G-Line	449.216	560	1636.94	3395.38
W-line	632.489	923	1823.34	1213.88

MAPE Values	Multiple Regression Internal Forecast	Winter's Method Internal Forecast	Multiple Regression External Forecast (Drought Scenario)	Winter's Method External Forecast
J-Line	11.0620	11	34.1236	57.3597
G-Line	11.2727	11	59.3978	50.3725
W-line	10.6468	11	35.3824	20.0018

[12] Delurgio, 1998. Pp. 43-55.

B-5 Descriptive statistics of weather conditions for use in Forecasting

The table below shows a sampling of the assumed values for weekly temperature and Precipitation used for forecasting ridership on Unitrans buses. Full table of values can be requested from HAT Consulting.

Week Starting	Avg Temp	max temp	min temp	Avg Precip	min precip	max precip
5/21/17	83	86.5714	79.8571	0.04667	0	0.14
5/28/17	89.8095	95.2857	84.4286	0	0	0
6/4/17	91.2380	93.5714	87.8571	0	0	0
6/11/17	85.9524	91.7143	79.1429	0	0	0
6/18/17	91.4762	93.8571	87.8571	0.0033	0	0.01
⋮	⋮	⋮	⋮	⋮	⋮	⋮
4/22/18	79.5714	85.8571	73.5714	0	0	0
4/29/18	78.0714	85.71428	74.5714	0.0375	0	0.13
5/6/18	80.6071	89.7143	73.1429	0	0	0
5/13/18	78.8036	84.2857	74.5	0.15	0	0.6

B-6 Average Forecasted values May 2017 – May 2018

Forecasted Average Weekly Ridership			
Line/Year	**El Nino**	**Normal**	**Drought**
J-Line	7,385	7,175	7,082
G-Line	6,458	6,103	5,824
W-line	9,451	9,266	9,183
Total	**23,294**	**22,544**	**22,089**

B-7 Weeks of the year with peak Ridership

A threshold of 26% increase in ridership, which is the average of the correlation of precipitation on ridership, was used to determine peak weeks. Commonality of peak periods were found among all forecasted scenarios and those periods are presented in the table below:

Forecasted Peak Demand Weeks					
2	3	5	6	7	9
10	11	15	17	18	41
42	43	44	45	47	49

References

1. DeLurgo, Stephen A. 1998. *Forecasting Principles and Applications*. Boston: Irwin/McGraw Hill.
2. Keller, G. (2012). Statistics for management and economics 10-th Ed. Mason, OH: Cengage Learning.
3. Palmere, Anthony. September 7, 2016. Unitrans General Manager's Report Fiscal Year 2015-16.
4. https://www.wildlife.ca.gov/conservation/marine/el-nino
5. https://www.nationalgeographic.org/encyclopedia/la-nina/
6. http://unitrans.ucdavis.edu/about/