USING GLOBAL POSITIONING SYSTEMS TO IMPROVE SCHOOL BUS ROUTING AND SCHEDULING

Principal Author:

Tori D. Rhoulac
Graduate Student
Institute for Transportation Research and Education
North Carolina State University
Raleigh, NC 27695-8601
919.515.6456, office
919.515.7809, fax
tdrhoula@eos.ncsu.edu

Secondary Authors:

Dr. Nagui Rouphail
Professor
Department of Civil Engineering
North Carolina State University
Raleigh, NC 27695-7908
919.515.1154, office
919.515.7809, fax
rouphail@eos.ncsu.edu

Jeffrey C. Tsai
Director
Pupil Transportation Program
Institute for Transportation Research and Education
Raleigh, NC 27695-8601
919.515.7931, office
919.515.8898, fax
tsai@unity.ncsu.edu

July 24, 2000

TRB Paper No. 01-3050

Abstract

School bus routing and scheduling in North Carolina is completed using TIMS, the Transportation Information Management System, according to state legislative mandate. TIMS uses an optimization algorithm to generate "acceptable" route alternatives to best serve the transported students and minimize the required number of buses. This algorithm utilizes link travel times and speeds encoded in the geocode (digitized map) to determine the routes for each school bus in the state. The speeds and travel times currently used by TIMS have no mathematical basis and have not been calibrated for conditions in North Carolina. The encoded speeds are simply reductions of the posted speed limit to account for student stops and the bus's inability to reach the posted speed. There is a need for models that can accurately estimate link travel times and student loading times so that TIMS can provide better estimates of optimal routes.

Mobile Global Positioning Systems (GPS) equipment was used on-board several school bus routes throughout the state to collect the data necessary for model development and validation. Two student loading time models were developed: a general model, which can be used in an urban, suburban, or rural area and an urban-specific model, which involves fewer predictor variables. Link travel time models were developed according to the two most pertinent factors: link length and traffic control device type at the link's end. Model validation was successful and now more efficient school bus routes can be developed because of more accurate geocode speed/ travel time estimates.

Key Words: TIMS, geocode, link

Using Global Positioning Systems to Improve School Bus Routing and Scheduling

1. Introduction

1.1 Background: TIMS

As of 1992, school bus routing and scheduling in the state of North Carolina is performed using TIMS, the Transportation Information Management System, according to state legislation. There are four primary components to TIMS: the geocode, transportation database, student database, and optimization algorithm. The geocode is a digitized map of links and nodes accurately reflecting the streets and intersections of the road network in each district. This map provides the basis for the entire computerized system. Bus stop, bus run, and bus route information is contained in the transportation database. This database includes information such as the location of student stops and the sequence of stops that compose a bus run. Another critical element of the transportation database is the optimal path algorithm. TIMS calculates the shortest time path to complete a bus run or route, based on the encoded link speeds. This information is then paired with records from the student database including the name and grade of each child needing to be served by the bus and the total number of children requiring service.

The final component of TIMS is the optimization algorithm by which a number of “acceptable alternatives” of a route serving the transported students is generated. Three separate optimization phases comprise the general heading of optimization. The first is stop optimization whereby the optimal locations of bus stops are generated using data from the student and transportation databases. Run optimization is the second step that assigns bus stops to bus runs in the most efficient, logical manner. A bus run is the unidirectional trip from student pick-ups to a school in the AM period and from that school to student drop-offs in the PM period. North Carolina often uses the “tiering” system of bus runs, which introduces the concept of a route. Route optimization allows for staggering school beginning and ending times so that more than one bus run can be completed by one bus, forming the total route for that bus (ITRE 1981). A route, then, may have several tiers, each tier corresponding to a specific bus run and school. This is the overall goal of TIMS- to produce routes that best serve the necessary students with the largest savings in cost and time. This task of finding the optimal route, with constraints such as the arrival times for the bus at each school, is a non-linear programming application (Graham 1984).

Link travel times and speeds are the primary input into this school bus routing optimization algorithm. The GIS data require speeds for each link along a route. The optimal path algorithm requires travel time for each link in order to find the shortest path according to time. Currently, travel time estimates in TIMS are static and based solely on the premise that the school bus is not able to travel at the posted speed limit because of student stops, turn movements, and other limitations. No algorithms exist to accurately estimate travel time for a link. This has led to the production of inefficient, sometimes infeasible, bus routes because of speed and travel time estimates that do not closely reflect actual conditions. Also, a “double-counting” bias is introduced between the inflation of link travel time considering student stops and the supplemental addition to travel time between stops suggested by the software that estimates the addition to travel time needed because of the number of students loading/ unloading at each stop.

1.2 Problem, Objectives, Scope, and Methodology:

The speeds currently used in the TIMS geocode are not only static with no mathematical foundation, but they have never been calibrated or validated for the specific conditions in North Carolina. Field tests completed in two North Carolina counties in the Fall of 1999 have proven that the speed estimates currently used in TIMS are inaccurate and may not reflect actual conditions on the road network. There is a need for a model that will predict accurate link speed and travel time estimates for TIMS data managers to input in their respective district geocodes. This model should be one that will estimate speed based on parameters such as the number of lanes in the direction of travel and a traffic congestion level, whose values can be modified to reflect future changes in traffic and roadway conditions.

Another component of TIMS that has not been calibrated for conditions in North Carolina is the time added to the bus run time to account for student loading and unloading. The current software, LOADTIME, suggests a fixed addition to the average travel time between student stops, based on the number of students loading or unloading at a stop. These values are suspected by pupil transportation officials to grossly overestimate the time required for actual student loading and unloading.

In response to the determined and expressed needs for better link speed and student loading time estimates, the research objective was established- **to develop an accurate, cost-effective, and quick means of measuring bus link travel times and student loading times so that empirical models could be developed.** This objective was completed using data collected with mobile Global Positioning Systems (GPS) equipment on-board several school bus routes. The empirical models were to provide methods to:

- Estimate student loading and unloading times.
- Estimate the average travel time required for a school bus to traverse links of various lengths with differing traffic and geometric parameters, independent of student stops.

The scope of this research is, then, to provide link travel time and student loading/ unloading time estimates based on regular travel events, such as turn movements, encountering traffic control devices, and delays due to traffic congestion.

A Global Positioning System (GPS) device was used to collect all data in this research. The GPS receiver records the position of the school bus at each second in time, giving accurate length and speed data for each link. Details of the equipment are provided in Section 2.2.

Field data collection was completed in three area types (urban, suburban, and rural) in order to model average conditions for the state. The primary differences between the area types are speed limits, link lengths, and link traffic control devices. Urban areas have shorter link lengths and a more variable range of speed limits than a suburban or rural area. Rural areas typically have higher speed limits, long link lengths and few traffic control devices.

1.3 Paper Organization

This paper is organized as follows: Section 2 covers the data collection and reduction and Sections 3 and 4 address the student loading and link travel times, respectively. Section 5 presents the research summary and conclusions, along with recommendations.

2. Data Collection and Reduction

2.1 Experiment Design and Objectives

Pupil transportation directors from three school systems selected the specific routes to be covered in each area. The selection process was based on area type (rural, suburban, or urban) and student behavior. Determining the sample size involved calculating the number of links needed to reflect the variety of conditions throughout the state. There were six data fields deemed important for data collection. These were speed limit, number of lanes in direction of travel, traffic control device type, turn movements, link length, and constricting geometric and pavement characteristics. Considering the possible levels for each of these and using the mathematics of combinations, the sample size was calculated to be 1620. Data on at least 1620 links was to be collected to adequately represent the number of possible treatments.

All measurable, contributing factors to the bus travel speed were recorded for each link of the bus route. These factors were determined to be:

- link length
- link travel time
- causes of deceleration
- speed limit
- number of lanes in direction of travel
- turn movement and traffic control device at link end
- pavement condition (paved, poor, or unpaved)
- constricting geometric conditions (narrow lanes, speed-restricting horizontal curves, or no lateral clearance)
- number of student stops
- number of students loading/ unloading at each stop
- loading time at student stop
- type of stop (address or corner)
- grade level of students loading/ unloading (elementary, middle, or high)

2.2 Data Collection Equipment

The application of Global Positioning Systems was the key component of the data collection process. The equipment used consists of three primary components: a data logger, an integrated GPS/ beacon antenna, and the GPS receiver. Figure 1 shows the equipment, with the receiver stored in a backpack. The data logger is a hand-held computer with a full alphanumeric keyboard and 8-line, 20-character display screen. The antenna integrates the GPS and beacon antennas into a single unit. This device intercepts the signals from the orbiting satellites and sends them to the receiver where the position and time information is stored. A position is recorded for each second of time, as established by the atomic clock.

Figure 1. GPS Equipment- Data Logger, Antenna, and GPS Receiver in Backpack

Collected data are called features. Features can be point, line, or area features. (Trimble 1998) For this application, line and point features are needed. Line features record data on the links and point features capture data occurring at specific locations such as student stops and causes of delay. The point features captured during data collection were student stops, stops for railroad crossings, deceleration due to traffic congestion, and deceleration due to a turn around movement.

The GPS receiver collected the remaining data internally. These data were the link length, link travel time, turn movement made at the end of each link- through, right, or left, and the duration of each student stop.

2.3 Data Reduction

Approximately 2936 links were covered in data collection. These data had to be post processed because of selective availability. The Department of Defense created this intentional scrambling of satellite signals when GPS came into widespread use, but differential correction removes this error and provides accurate position reports. Selective availability has since been alleviated, however. Following differential correction, data were exported into a spreadsheet for analysis.

3. Student Loading Time

3.1 Model Development

To establish the need for a student loading time model, six routes were selected, two from each area type, and the actual loading times observed in the field were compared with the TIMS suggested loading times in LOADTIME. A set time penalty is taken from a chart, based on the number of students loading or unloading at each stop. Each of the individual stop loading times is added over the entire route. Table 1 displays the results of the TIMS loading time comparison and the TIMS speed comparison, discussed in Section 4.1.

The observed loading time for Bertie1 is similar to the TIMS loading time suggestion. The remaining five routes, however, do reveal vast differences between the actual loading and TIMS suggested times. In general, TIMS overestimates the required time for student loading. These results prove that a model is needed to provide more accurate loading time estimates.

The factors that may have an impact on student loading time are: number of stops per link, total number of students loading per link, the interaction between these stops and students, the student’s grade level, and the area type. Currently, TIMS suggests loading times based solely on the number of students. Figure 2 is a graph of the number of students versus observed loading times in the field. This plot shows a definite linear

Table 1. TIMS Loading Time and Speed Comparisons

Route ID	**Total Stops on Route**	**Total Students on Route**	**Observed Loading Time (sec)**	**TIMS Loading Time (sec)**	**Loading Time Difference (sec)**	**Average GPS Speed (mph)**	**Average TIMS Speed (mph)**	**Speed Difference (mph)**
Bertie 1	16	29	442	444	-2	---	---	---
Bertie 6	21	35	422	816	-394	---	---	---
Cary 2	31	103	812	1368	-556	---	---	---
Cary 4	14	15	209	480	-271	---	---	---
Nral 3	7	27	166	348	-182	---	---	---
Nral 5	15	82	477	786	-309	---	---	---
Bertie 5	---	---	---	---	---	30	38	-8
Bertie 10	---	---	---	---	---	35	45	-10
Cary 1	---	---	---	---	---	22	16	6
Cary 3	---	---	---	---	---	24	26	-2
Nral 2	---	---	---	---	---	28	38	-10
Nral 5	---	---	---	---	---	27	28	-1

pattern as the number of students increases, but there is large variability in loading time for the lower range of students loading. Observing the characteristics of the student stops that collectively establish a linear pattern reveals that each of these points is from urban data. Figure 3 shows a graph of the number of stops versus observed loading times. Although there is large variability in loading time on links with one and two stops, a pattern develops as the number of stops per link increases, indicating the importance of stops as a variable in non-urban areas.

The linear pattern in the urban data suggests that the number of students, N_u, is sufficient to estimate the average student loading time, T_{LU}, in urban areas. Performing regression analysis on the urban data yields Equation 1, with output in seconds and an R^2 of 0.79 and estimated standard deviation of 7.7 seconds.

$$T_{LU} = 11.31 + 2.93 N_u \qquad (1)$$

Loading time can be modeled differently for urban students because of bus stop location. Typically, urban bus stops are at corners, as opposed to specific addresses. The time required for students to load from a corner is relatively shorter than loading from a specific address, which is common in suburban and rural areas. Also, in rural areas where there are no safe walkways, stops are more frequent along a single link because the bus must stop at each student's driveway. The importance of stops as a model variable increases then, when the area type is rural or suburban. These differences account for the predictability of urban student loading time by solely the number of students loading and the necessity of other predictor variables in suburban and rural areas.

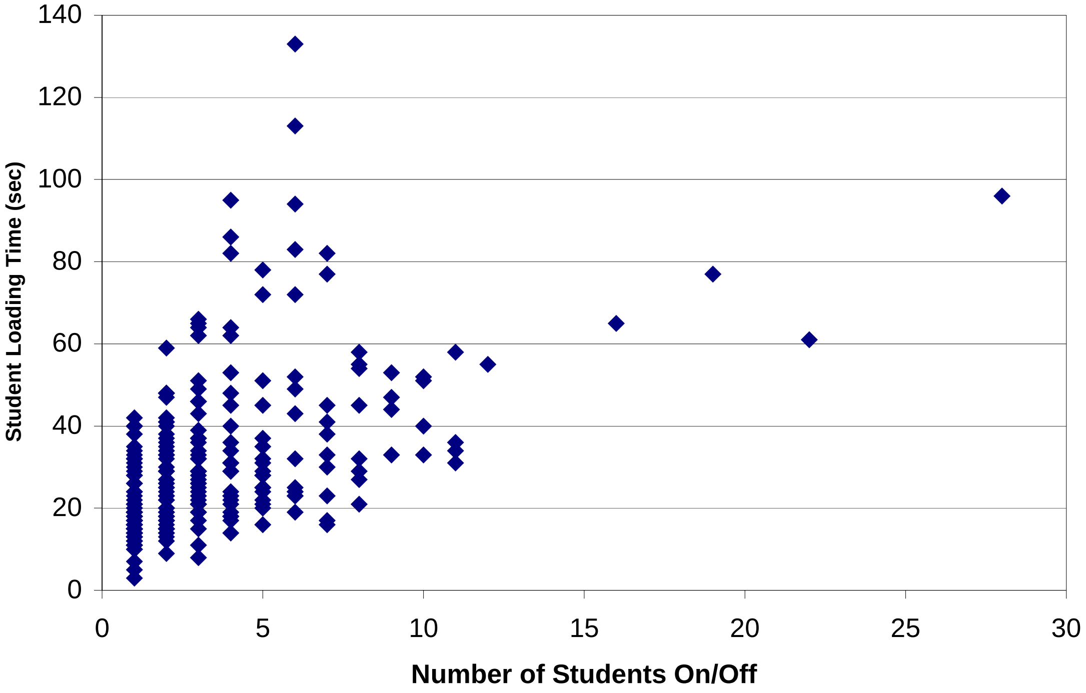

Figure 2. Student Loading Time versus Number of Students

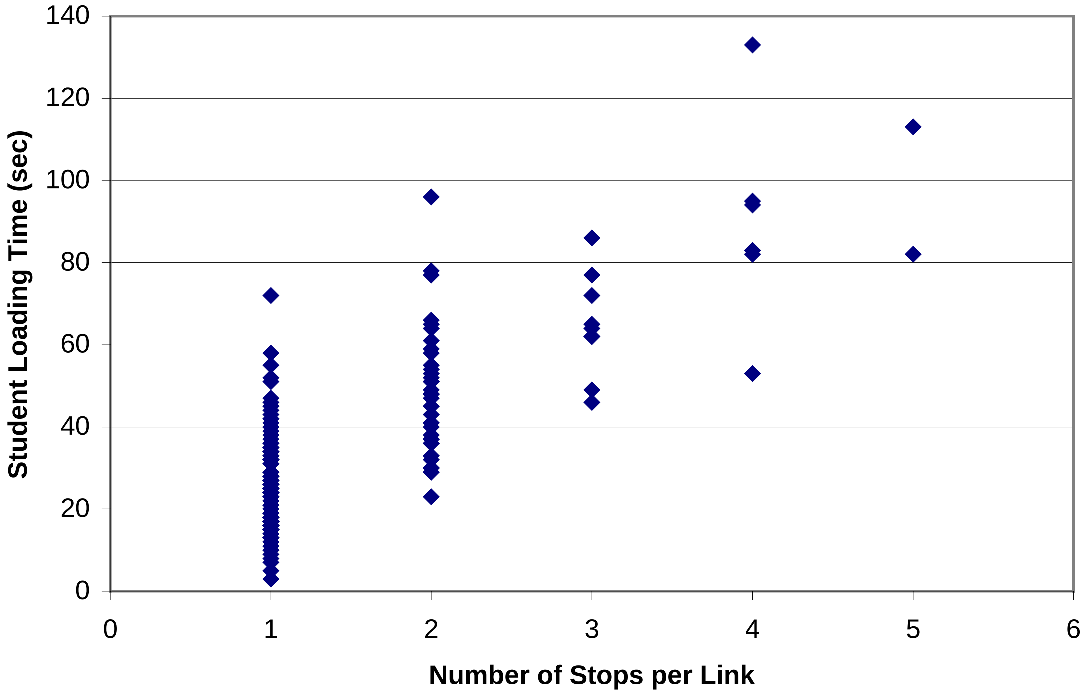

Figure 3. Student Loading Time versus Number of Stops

Some routes, however, may go through a mixture of urban, suburban and rural conditions, making a general model necessary that can be used regardless of the area type. The second analysis will include the suburban and rural data, and the urban data in an attempt to generate a single, general model.

Regression analysis on the combined data set, using the three most pertinent predictor variables- students, N_u, stops, N_o, and the interaction, yields Equation 2 with an output in seconds and an R^2 of 0.7 and estimated standard deviation of 9.8 seconds.

$$T_{LG} = 3.56 + 0.74N_u + 13.77N_o + 1.47(N_u \times N_o) \tag{2}$$

To assess the impact of grade level, additional regression analyses involving only elementary, middle, and high school data, respectively, will be performed in the three-variable equation format. The analysis involving elementary students presents findings that are inconsistent with the original conclusion that a direct relationship exists between loading time and the number of stops and students on a link and the interaction. In this model, only the number of stops per link variable is statistically significant. Accordingly, a time penalty of 20.2 seconds will be applied to each stop along a route, regardless of the number of students loading or unloading. An additional model may be necessary treating elementary students separately, but analysis of the middle and high school cases must first be completed.

Based on the middle school model, a 17.6-second penalty would be accessed to the route travel time for each stop and an additional 2.9 seconds would be added for each student loading or unloading. Therefore, the minimum penalty for each stop along a route carrying middle school students would be 20.5 seconds. This is only a 0.3-second difference between the time penalty suggested by the elementary school model based only on the number of stops per link.

The high school model does not follow a similar pattern, however. There are no statistically significant coefficients reported by this model. A separate model for high school students, therefore, could not be justified.

Based on these results, grade level does not affect the student loading time in such a way that can be effectively modeled. The three-variable model resulting from the comprehensive analysis assesses at least a 19.4 second penalty for each stop on a route, regardless of grade level. The middle and elementary models suggested minimum penalties that differed by only about 1 second. The high school data did not yield substantive values. The effect of grade level can be omitted from the model.

3.2 Model Validation

The F-test is the method of statistical evaluation used to validate the model results. For this application, standard 95% confidence limits were used, corresponding to an α of 0.05. A second phase of the F-test was completed to account for any county effects. These calculations were completed by replacing the d_i^2 in the equation used to calculate the F-statistic with $(d_i - d)^2$. Equation 3 shows how this F-statistic, denoted by F_c, is calculated.

$$F_c = \frac{\frac{1}{N}\sum_{i=1}^{N}(d_i - d)^2}{s^2} \qquad (3)$$

Data from 31 student stops on a middle and high school route were used in model validation. The specific loading times are compared with the suggested loading times of the newly created model. According to F-test results, the model-predicted loading times can be considered statistically equivalent to the actual loading times with 95% confidence. The variance of the model is 95.16 seconds squared. Not considering county effects, the variance of the model loading times is 176.43. Dividing the variance of the model predictions by the model variance results in an F-statistic of 1.85. The second set of calculations, accounting for county effects, resulted in an F-statistic of 1.83. According to F tables, the F value corresponding to 31 degrees of freedom in the numerator and 30 degrees of freedom in the denominator at an α of 0.05 is 1.84 (Mendenhall 1996).

The theory of the F-test establishes that a null hypothesis cannot be rejected if the F-statistic is less than the table value. The null hypothesis for this application is that the model predictions variance equals the model variance, allowing for the inference that the two populations are also equal. This hypothesis can be rejected at a 95% confidence level because the F-statistic of 1.85 exceeds the table value of 1.84. Accounting for county effects, however, the null hypothesis cannot be rejected because the F-statistic of 1.83 is less than the table value of 1.84.

Validation of the student loading time model has proven that this general model does yield acceptable results when predicting average student loading times. No data were collected for model validation in an urban area, so the urban-specific model cannot be validated. The mathematical student loading time models can be reviewed in Equations 1 and 2 above.

4. Link Travel Time

4.1 Model Development

The fundamental unit of analysis is the link. All field data were collected by link and will be analyzed based on the individual link.

To establish the need for a link travel time model, two routes from each of the three area types were selected and the observed field speeds were compared with the TIMS geocode speeds for those specific routes. For ease of comparison, the individual link speeds were averaged for each route. The average route speeds and comparison results are presented in Table 1, found in Section 3.1. While routes Nral 5 and Cary 3 differ in average speed by less than 5 miles per hour, the remaining routes do reveal an average speed difference large enough to support the necessity of empirical model development.

The first step in developing the model for estimation of link travel times is determination of the independent and dependent variables. Originally, link travel speed was the desired dependent variable on which regression analysis would be performed. For purposes of continuity with the student loading time model, however, link travel time was deemed to be a better choice. With link travel time as the dependent variable, both models now estimate time parameters in seconds. The independent variables that will be included in the model are: free-flow travel time (computed by dividing the speed limit into the link length), number of lanes in direction of travel, traffic control device, turn movement, stops for congestion (including stops at railroad crossings), restriction codes, and the interactions of the congestion, turn movement, and restrictions variables with free-flow travel time. The non-numeric variables for turn movements and restrictions are converted to dummy, numeric variables for the regression analysis. The symbols used to represent each variable in the regression are given below.

- link travel time, in seconds (time)
- free-flow travel time, in seconds (fftime)
- number of lanes in direction of travel (lanes)
- stops for congestion (cong)
- turn movement dummy one (tm1)- equals 1 if bus makes a left turn, equals 0 for through and right turn movements
- turn movement dummy two (tm2)- equals 1 if bus makes a right turn, equals 0 for through and left turn movements
- restriction code dummy one (rc1)- equals 1 if bus encounters speed-restricting narrow lanes, equals 0 if not
- restriction code dummy two (rc2)- equals 1 if bus encounters speed-restricting horizontal curvature, equals 0 if not
- restriction code dummy three (rc3)- equals 1 if bus travels on unpaved link, equals 0 if not
- congestion and fftime interaction (congT)
- left turn and fftime interaction (tm1T)

- right turn and fftime interaction (tm2T)
- narrow lanes and fftime interaction (rc1T)
- horizontal curvature and fftime interaction (rc2T)
- unpaved roads and fftime interaction (rc3T)

Traffic volumes are also significant in estimating link travel times, but were not included because of resource constraints. Volume data are not available for a majority of the links on which school buses travel, including residential streets. The resources required to obtain these volumes would have exceeded the time and budget allotted for this research. Therefore, stops for congestion was included as a variable to indirectly account for link volumes by measuring instances of bus interaction with other vehicles.

The two factors that will impact link travel time most are link length and traffic control device type at the link end. These are not included directly as independent variables because the composite data set has been separated into six categories based on length and traffic control device type. Preliminary analysis illustrated that links less than 2640 feet (one-half mile) behave similarly as do links greater than one-half mile. Two length categories resulted: short, denoted by SS identifying links less than 2640 feet, and long, denoted by LL identifying links greater than 2640 feet in length. There are three traffic control device categories: NC (no control), TP (stop sign, including yield signs), and SG (traffic signals). Yield and stop signs are grouped because as school buses encounter yield signs, they are observed to stop as a safety precaution. No control is necessary because the bus may pass a minor street with two-way stop control, but the bus itself encounters no traffic control.

Table 2 displays the regression analysis results for each of the six categories. The number of independent variables is noted because independent variables with a zero sum were removed from the analysis. Therefore, all categories do not have the same number of independent variables of the fourteen possible.

Four models have low R^2 values and/ or high variances. The number of different situations that are possible on long links explain the high variance of the TP-LL model. The SG-LL model involved only ten links and four independent variables: free-flow travel time (fftime), lanes in direction of travel (lanes), right turns (tm2), and the right turn-time interaction (tm2T). The remaining ten predictor variables had zero values.

There is no practical use for this model. Right turns are significant only because there are not enough data to fully examine the variable relationships for signal-controlled links longer than 2640 feet. These ten data points should be combined with the SG-SS model for a composite, signal-controlled model category.

Table 2. Link Travel Time Regression Results

Model	**Number of links**	**R^2**	**Variance (sec^2)**	**Number of Independent Variables**
NC-SS	1494	0.61	65.2	14
NC-LL	125	0.95	169.3	14
TP-SS	204	0.39	195.7	14
TP-LL	23	0.97	649.9	13
SG-SS	552	0.29	532.4	10
SG-LL	10	0.95	32.7	4

The composite SG model is comprised of 562 links with an R^2 and variance of 0.32 and 526.9, respectively. The significant variables are free-flow travel time, stops for congestion, left and right turns (tm1 and tm2), and narrow lanes (rc1).

Both the SG and TP-SS models have R^2 values less than 0.6, which is not desired. Residual analysis, however, indicates neither that another model form would be more appropriate nor that standard forms of error are present. Other statistical error analysis methods were considered but deemed inapplicable for this application. Time series analysis, for example, may have helped to explain more of the variance, but because there are five models used to represent the data, each involving links of different characteristics, the links were not analyzed sequentially. Therefore, time series analysis could not be utilized. No changes will be made to these models. Table 3 summarizes all five models, giving the number of links involved, the R^2 and variance, and the coefficients for each of the variables.

Table 3. Link Travel Time Models Summary

	Number of links	Model Intercept	Variable Coefficients						R^2	Variance (sec^2)
			fftime	**cong**	**tm1**	**tm2**	**rc1**	**rc3**		
NC-SS	1494	1.86	1.11	10.12	5.30	---	---	---	0.61	70.2
NC-LL	125	---	1.35	---	---	---	---	---	0.93	217.0
TP-SS	204	14.25	1.01	17.01	---	---	---	---	0.36	189.2
TP-LL	23	---	1.21	---	---	---	---	45.09	0.93	715.8
SG	553	10.51	1.13	17.85	13.89	6.56	47.41	---	0.31	527.8

4.2 Model Validation

Data from 173 links were used in model validation. The statistical procedure used to evaluate these data was the F-test. An overview of the F-test procedure is presented in Section 3.2, Student Loading Time Model Validation.

The validation data were categorized according to link length and traffic control device to correspond directly with the five model types. For each model type, the F-statistic was computed with and without accounting for county effects. F_c denotes the F-statistic accounting for county effects. All results and other pertinent values involved in the validation process are shown in Table 4. Model validation was successful in that the equal population variance hypothesis is not rejected for any model. The F-statistic is less than the F_{table} value for all models. The F-test is passed for the TP-SS model because the F_c is lower than the F_{table} value. These models will be used throughout North Carolina, so accounting for county effects is recommended as this decreases the sample variance by considering the average observed travel time for the specific county under consideration. Therefore, each of the link travel time models will not be rejected and is considered an effective means of estimating link travel time for school bus operation.

Table 4. Link Travel Time Model Validation Results

Model	**N**	**Sample Variance (D)**	**Sample Variance** w/ County Effects **(D')**	**Difference (D - D')**	**Model Variance (s^2)**	**F**	**F_c**	**F_{table}**	**Reject Model**
NC-SS	112	38.43	36.77	1.66	65.55	0.59	0.56	1.42	no
NC-LL	12	282.02	274.32	7.70	216.98	1.30	1.26	2.79	no
TP-SS	24	455.16	346.60	108.56	189.21	2.41	1.83	2.01	no
TP-LL	5	2257.59	1067.37	1190.22	715.818	3.15	1.49	6.26	no
SG	20	797.24	778.13	19.11	527.78	1.51	1.47	2.16	no

Of particular note are the F-statistics for the NC-SS model being less than one, which is a rare occurrence in analysis of variance. The reason that the sample variance (numerator) is less than the model variance (denominator) is that the validation links are better data, in general, than the model development data. Performing regression analysis on the model development data yielded an R^2 and variance of 0.61 and 70.2, respectively. Regression of the validation data yields an R^2 of 0.81 and variance of 37.1. The sample variance is therefore less than the model variance because a linear model better fits the validation data than the original data collected in the field.

5. Summary, Conclusions, and Recommendations

5.1 Summary

The results of this research are mathematical models to supplement TIMS, the Transportation Information Management System software used for school bus routing in the state of North Carolina. TIMS uses optimization methods to produce the most efficient, effective school bus routes and requires as input, the travel speed for each link. The empirical models developed will provide more accurate travel time estimates in order to produce more optimal routes. The two most important contributors to the travel time of a school bus on a link are 1) traffic and roadway conditions and 2) student loading. An accurate, cost-effective, and quick means of measuring bus link travel times, student loading times, and the associated factors was desired for data collection. Mobile GPS equipment was used on-board several school bus routes to collect this data for model development.

Using linear regression, two categories of empirical models were developed to estimate link travel times and student loading times. The primary variables used to estimate travel time in the five, link-based models are link length, posted speed limit, and traffic control device type. Two student loading time models resulted based on the area type (urban, suburban, or rural) in which the route is located. Student loading time is directly related to the number of students loading in urban areas where most student stops are at corners, not specific addresses. An average link length of approximately 700 feet and the frequent occurrence of traffic signals define an urban area. The individual stop becomes more important in suburban and rural areas where stops are more frequent. A second, general model was developed to estimate student loading time based on the total number of students loading on a route, the total number of stops on the route, and the stops-students interaction. The general model can be applied in urban, suburban, or rural areas. This is necessary because a single route may cover more than one area type.

Both the student loading and link travel time models were successfully validated using F-tests. The models have proven to estimate student loading and link travel times better than the current TIMS methodologies. Further tests have also verified the utility of the models. TIMS reported the route time of a randomly selected sample route as 67 minutes. The driver of that route reported an actual route time of 105 minutes. Application of the link travel time and student loading models to this route led to a route time of 107 minutes, only a 2 minute difference from the driver-reported route time.

The statistical testing of the models is presented in Sections 3.2 and 4.2. More detailed statistical analyses were not completed as this was beyond the original research scope and because of the nature of the data, which were not analyzed as time series data.

5.2 Conclusions

1. *Considering the data available in TIMS, school bus travel time cannot be effectively modeled in areas with high traffic signal frequency.*

Although the signal controlled link travel time model yielded favorable results in model validation, the R^2 and variance of the model suggest that the travel time on signal controlled links cannot be effectively modeled without additional information that is not contained in TIMS. Traffic signals occur most frequently in urban areas where many atypical events occur to impact the travel time of a bus. These extraneous events make accurately estimating link travel time on signalized links difficult. Also, travel time on links with signals is dependent on the occurrence of green. Encountering a signal in the green phase involves a much shorter travel time than having to stop on a link with a signal in the red phase. Involving signal timing data would better the reliability of a link travel time model on signalized links, but this level of detail is beyond the scope of this research.

2. *The most significant factors affecting school bus link travel times are:*

- *Posted speed limit*
- *Link length*
- *Traffic control device at link end*
- *Pavement quality*
- *Lane width*
- *Traffic congestion*

Only the posted speed limit is utilized currently in TIMS to estimate the link travel speeds and associated route time. The results of this research indicate that posted speed limit alone cannot accurately predict school bus speed. Other factors such as link length, traffic control device type, lane width, level of traffic congestion, and whether or not the link is paved are critical.

3. *Student loading time is highly dependent on whether the stop is at an address or corner.*

In rural and suburban areas, many stops are address stops, meaning that the bus stops at a specific residence for loading or unloading. Address stops are necessary when there is no "safe" walkway for the resident(s) of that address to meet with other students at a corner bus stop. More stops are involved when buses are required to stop at specific addresses rather than corners. The time associated with the individual stop is therefore more important for routes with frequent address stops. When the majority of stops are at corners, the number of students loading or unloading can alone determine an accurate loading time estimate. Type of stop- address or corner- therefore, has significant impact on the required loading time.

5.3. Recommendations

The scope of this study should be expanded in future research to better calibrate the models for the range of possible conditions in the school bus travel environment throughout North Carolina. Transferability is vital as the models will be implemented statewide. Model development data were collected in only central and eastern North Carolina. Western North Carolina was not included because of budgetary constraints, but this mountainous region may largely impact the results. Further study should incorporate data collection in at least one western North Carolina County. More data would also help to increase statistical confidence in model variables.

Continuing research on the application of Global Positioning Systems and other ITS technologies to pupil transportation operations and safety is imperative. GPS proved to be an essential resource in the process of bettering the routing and scheduling process. Continuation of this initiative has great potential to be increasingly beneficial to the pupil transportation and transportation engineering communities.

Acknowledgements

I must first thank God and Dr. Nagui Rouphail (Civil Engineering, NCSU), Dr. John Stone (Civil Engineering, NCSU), Dr. John Monahan (Statistics, NCSU), and Jeff Tsai, Director of the Pupil Transportation Program at the Institute for Transportation Research and Education for their priceless contributions, advice and guidance. A special thank you to Derek Graham and the North Carolina Department of Public Instruction for funding this research. I offer my sincere appreciation for the cooperation and assistance of the entire Pupil Transportation staff at ITRE, Mr. Wyatt Currin and the Wake County Transportation Staff, Mr. Phil Bragg, Mrs. Vernice Murphy and the Bertie County Transportation Staff, and Ms. Catherine Ward, formerly of the GIS group at ITRE.

References

(1) Graham, Derek S. Computer-Assisted School Bus Scheduling. Masters Thesis, North Carolina State University, 1984.

(2) Mendenhall, William and Terry Sincich. A Second Course in Statistics. Upper Saddle River, New Jersey: Prentice-Hall, Inc., 1996.

(3) Trimble Navigation, Pro XR/XRS with Pathfinder Office: Classroom Manual. Sunnyvale, California: Trimble Navigation Limited, 1998.

(4) “Optimized School Bus Routing and Scheduling for North Carolina.” Research Triangle Park, North Carolina: Institute for Transportation Research and Education (ITRE), 1981.